The Dissertation Committee for Valerie Jean Bradford Certifies that this is the approved version of the following dissertation:

Aromatic Donor-Acceptor Interactions: Bridging Abiotic and Peptide Folding

Committee:

Brent L. Iverson, Supervisor

Eric V. Anslyn

Michael J. Krische

David W. Hoffman

Laura J. Suggs

Aromatic Donor-Acceptor Interactions: Bridging Abiotic and Peptide Folding

by

Valerie Jean Bradford, B.S.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May 2008

Dedication

To Lafe and my family.

Acknowledgements

I would like to that Professor Brent Iverson for his support and patience. His positive thinking and the phrase “It will work now that you’ve given up” helped me to keep motivated and continue on many projects that I would have otherwise given up on. I want to thank everyone at UT who has helped me, especially the members of the Willson, Anslyn, and Sessler labs. I would especially like to thank members of the Iverson group for all there help and support through the years. Karl Griswold and Navin Varadarajan for helping me with molecular biology and letting me share their benches (but not their pipettmen!). Greg Gabriel, thanks for sharing my pain with the unpredictable gelling compounds. Joe Reczek, thanks for giving me advice whether in chemistry, life, or my career. Thanks to Chelsea Martinez for all her help in editing this dissertation and for not letting me be the only female in the group. Stevan Samuel was always there to provide entertainment. I will certainly miss his company and irrational statements, although I’m sure I will continue to see them on Facebook. I’d like to thank Garen Holman for all his help in my last year here both in an out of lab. Good luck to all of you in your careers!

I’d like to thank my Mom and sister for their love, visits, and listening to me when I needed to vent. Most of all, I would like to thank my husband, Lafe, for all his love and support. He always believed in me, even when I didn’t. Even in phone calls from Korea, the frozen tundra of New York, and Iraq, a “you’ll be ok” or “I know you can do it” helped lift my spirit and make it through some very tough times.

Aromatic Donor-Acceptor Interactions: Bridging Abiotic and Peptide Folding

Publication No._____

Valerie Jean Bradford, Ph.D.

The University of Texas at Austin, 2008

Supervisor: Brent L. Iverson

Aromatic donor-acceptor interactions have been utilized by the Iverson group in the development of abiotic molecules, called aedamers, that achieve new folding motifs, intermolecular association in heteroduplexes, and new material properties. These molecules exploit the interaction between the electron-rich 1,5-dialkoxynapthalene (DAN) and electron-deficient 1,4,5,8-naphthalenetetracarboxylic diimide (NDI) units in a face-centered stacking geometry in aqueous solution.

This dissertation describes the use of DAN-NDI interactions in the realm of peptides and proteins to expand the scope for applications of this interaction. This work specifically focuses on three areas of aromatic donor-acceptor interactions: achieving protein behavior with abiotic molecules, introducing the interaction into natural peptides, and utilizing the interaction in the intermolecular association of an abiotic molecule and a natural peptide.

Chapter 2 refines the model of aggregation of an amphiphilic aedamer, which forms a hydrogel upon heating. The aedamer behaves similarly to proteins called

amlyoids, which form fibrils and plaques *in vivo* which have been implicated in a variety of diseases, including Alzheimer's. Chapter 3 describes the synthesis of α-amino acids with DAN- and NDI-containing side chains. These amino acids can be used in a peptide model of β-hairpin secondary structure. The model system can determine whether aromatic donor-acceptor interactions are useful in stabilizing peptide and protein structure. Chapter 4 describes the study of the Anchored Periplasmic Expression System (APEx) for use in screening random peptide libraries. A random peptide library is used to determine the sequence of a natural peptide, potentially containing electron-rich aromatic residues, which could bind an NDI oligomer with high affinity for use as a protein expression tag. Chapter 5 describes work toward the use of cyclic NDI bisintercalators for binding both the major and minor grooves of a specific sequence of DNA simultaneously, in addition to the use of cyclic NDI and DAN molecules for the further study of NDI-DAN interactions in abiotic intermolecular assembiles.

Overall, this work has advanced the application of aromatic donor-acceptor interactions in peptides and should serve as a foundation for the future study of this interaction in protein folding and behavior in biological systems.

Table of Contents

List of Tables

List of Schemes

List of Figures

CHAPTER 1

CHAPTER 3

CHAPTER 4

CHAPTER 5

Chapter 1

Self-Assembly and Recognition with Aromatic-Aromatic Interaction

1.1 AROMATIC INTERACTIONS IN PROTEINS

Aromatic interactions have been recognized in structural biology for their important contributions to DNA stability, protein folding, and protein-ligand recognition. While aromatic interactions are not as abundant as other non-covalent interactions, such as hydrogen bonding, they are important nevertheless. For example, the overall frequency of aromatic residues in proteins is very low; tryptophan (Trp) has the lowest frequency of any amino acid residue at 1.32% overall, tyrosine (Tyr) and phenylalanine (Phe) appear at 3.25% and 3.91%, respectively. Although the overall frequency of aromatic residues is low, they are highly conserved among a variety of protein families. Tryptophan is the most conserved residue, while phenylalanine and tyrosine third most conserved. (Gazit 2002) Aromatic residues are involved in a variety of non-covalent interactions, including aromatic-aromatic, cation-π , X-H-π , and S-π interactions.

1.1.1 Aromatic-Aromatic Interactions

The importance of aromatic-aromatic interactions in proteins was brought to the forefront of structural biology research by the work of Burley and Petsko. Data mining of 34 protein structures from the Brookhaven Protein Data Bank (PDB) revealed that 60% of the aromatic residues were involved in aromatic-aromatic interactions. In addition, 80% of the aromatic-aromatic interactions in the proteins stabilized tertiary structure by linking secondary structure motifs together, while the remaining 20% stabilized quaternary structure. (Burley 1985)

There are four geometric conformations of aromatic-aromatic interactions commonly found in proteins: parallel displaced, T-shaped, parallel staggered, and herringbone, as shown in Figure 1.1. While Burley and Petsko concluded that the T-shaped conformation was the predominate geometry in proteins, an extended analysis of a larger sample of proteins suggested that the parallel-displaced geometry was the preferred conformation. (Burley 1985, McGaughey 1998)

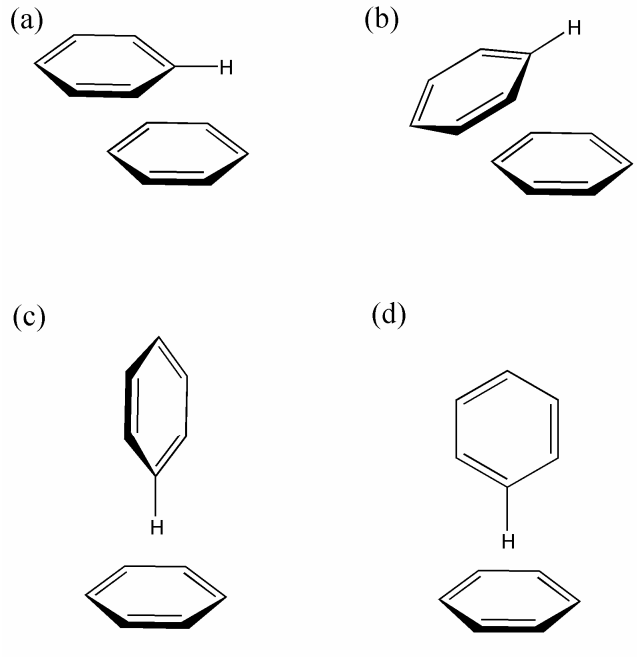


Figure 1.1 The four geometries of aromatic-aromatic interactions in proteins. (a) Parallel-displaced (b) parallel staggered (c) T-shaped (d) herringbone

The energetic contribution of aromatic-aromatic interactions was determined definitively by double-mutant cycles of the ribonuclease from *Bacillus amyloliquefaciens*, barnase. Barnase displays a T-shaped interaction between two tyrosine residues on a solvent-exposed helix. The double-mutant cycle of the tyrosine residues to alanine gave an interaction energy of -1.3 kcal/mol. A Phe-Phe pair

substitution showed an identical contribution to the stability of the protein. This interaction energy is similar to the contribution of hydrogen bonding of non-charged groups. (Serrano 1991)

1.1.2 Other Aromatic Interactions

In addition to classical aromatic-aromatic interactions, aromatic residues are involved in other non-covalent interactions, such as cation-π, X-H-π, and S-π interactions as shown in Figure 1.2. The biological implications of cation-π interactions have been reviewed extensively. (Ma 1997) The charged residues lysine, histidine, and arginine are most commonly found in cation-π interactions with Phe, Tyr, and Trp. The average interaction energy for cation-π interaction in proteins is between -0.4 and -2.4 kcal/mol. (Meyer 2003)

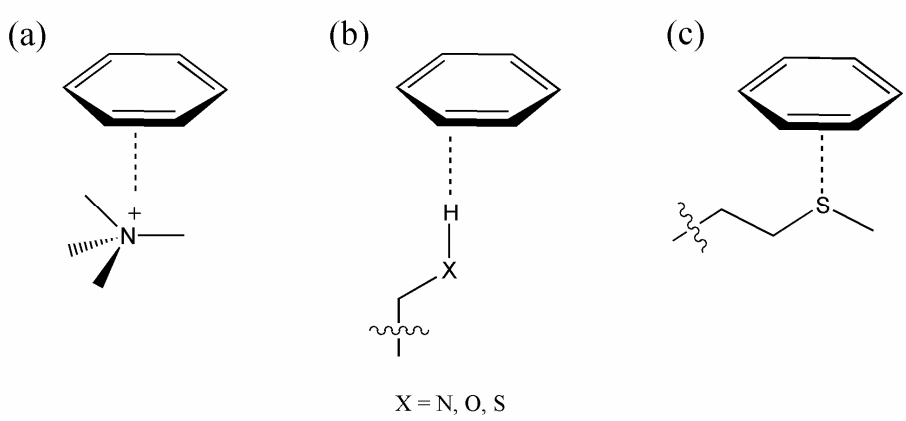


Figure 1.2 Other types of aromatic interactions (a) cation-π (b) X-H-π (c) S-π.

X-H-π interactions occur with N-H, O-H, and S-H groups in protein structures. This interaction is relatively common, with one out of every 11 aromatic residues acting as a π-acceptor for one of the previously mentioned groups. The most common aromatic residue in this type of interaction is tryptophan. (Steiner 2001) While the most obvious conformation for this interaction would be T-shaped, the preferred geometry in proteins,

as determined by a survey of the Cambridge Structure Database, is the hydrogen of X-H directly over a carbon atom in the ring, not over the center of the ring. (Malone 1997)

Sulfur-π interactions were first recognized in the 1970s. Eight proteins studied showed S-π interactions, including lysozyme from hen egg white. Lysozyme contains a chain of 7 alternating sulfur-aromatic residues as shown in Figure 1.3. (Morgan 1978, Diamond 1974) There are several important factors to be considered in S-π interactions including the polarizability of sulfur, the empty 3d orbitals on sulfur, and the well-known ability of sulfur-containing molecules to quench Tyr and Trp fluorescence. (Augenstine 1961) While the most common S-π interactions involve methionine or the free thiol of cysteine, a survey of immunoglobulin (Ig) protein structures showed a highly conserved disulfide-bonded cysteine-Trp interaction. (Freemont 1996) Empirical methods predicted the stabilization energy of this interaction to be 0.8 kcal/mol. (Nemethy 1981) Methionine is often found in enzyme binding pockets, where an aromatic substrate is sandwiched between an aromatic residue on one side of the ring and the sulfur of methionine on the other. (Meyer 2003)

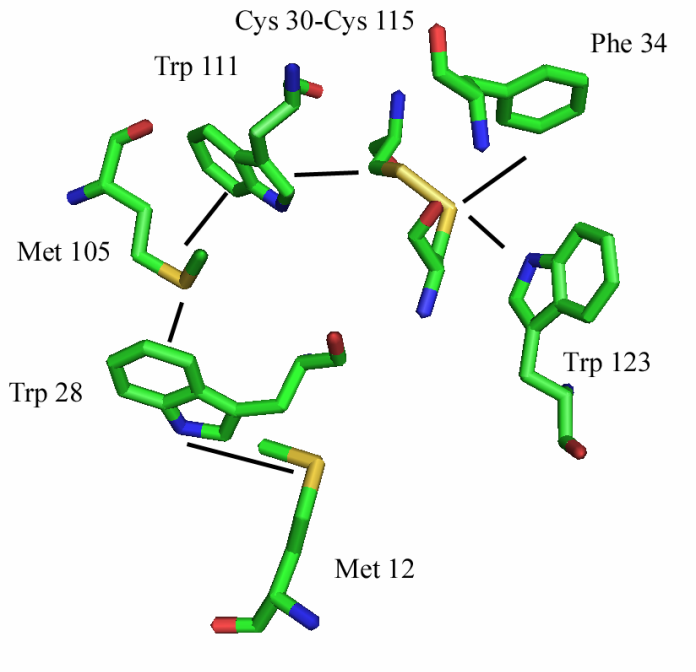


Figure 1.3 Alternating pattern of sulfur-containing and aromatic residues in lysozyme. (Diamond 1974, PDB: 6LYZ)

1.1.3 Aromatic Interactions in Protein-Protein and Protein-Ligand Interactions

In addition to mediating protein folding, aromatic interactions have also been shown to mediate protein-protein and protein-ligand interactions. A statistical analysis of a bacterial transmembrane database revealed an over-representation of an aromatic-XX-aromatic motif. This is analogous to the QXXS motif found in many transmembrane proteins which affords homodimerization by forming hydrogen bonds between the glutamine and serine residues. (Sal-Man 2007) The QXXS motif of the N-terminal TM domain of *E. coli* aspartate receptor (Tar-1) was replaced with the aromatic-aromatic interaction in a WXXW motif. The resulting mutant retained 75% of the wild type dimerization activity of Tar-1. In addition, the cholera toxin secretion protein EpsM was shown to have the WXXW motif naturally. Mutation of the aromatic-aromatic interaction to a AXXA sequence eliminated dimerization acivity of the EpsM protein and provided

strong evidence that the dimerization of transmembrane proteins with the aromatic-XX-aromatic motif is dependent on aromatic-aromatic interactions.

Aromatic-aromatic Interactions have been shown to stabilize co-factor binding in proteins. Analysis of hemoprotein structures in the PDB showed a large proportion of structures contained at least one Phe, Trp, Tyr, or His side chain involved in aromatic-aromatic interactions with the heme, as shown in bovine cytochrome c peroxidase in Figure 1.4. Tryptophan 191 displayed a parallel offset geometry, while tryptophan 51 showed a T-shaped geometry. Replacement of a Trp residue in a designed hemoprotein model system with an alanine residue reduced the percent of the folded population from 83% to 50% and by NMR, the heme was more tightly bound with the Trp residue. (Liu 1999)

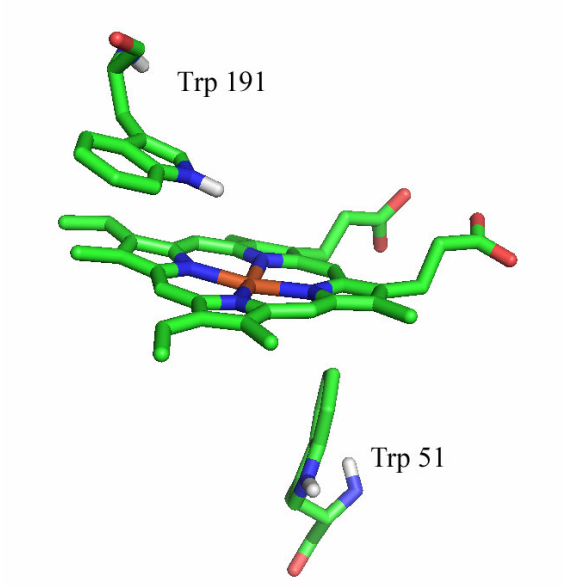


Figure 1.4 Heme binding pocket of bovine cytochrome c peroxidase showing the aromatic-aromatic interactions of Trp 191 and Trp 51 with heme. (Goodin 1993, PDB: 1CCA)

One of the more recent developments in aromatic-aromatic interactions has been their importance in carbohydrate recognition proteins. Carbohydrate binding proteins,

such as glycosidases, glycosyltransfreases, and lectins, have an abundance of aromatic residues in their binding pockets. A prime example of aromatic-aromatic interactions in carbohydrate recognition is human lysozyme, which recognizes homopolymers of β(1,4)-N-acetylglucosamine (GlcNAc). Lysozyme has a tyrosine residue parallel to the hydrophobic face of the carbohydrate residue and two tryptophan residues in proximity to the methyl and methylene group of the carbohydrate residue as shown in Figure 1.5. Replacement of the Tyr residue with Phe or Trp showed little effect on binding while substitution with leucine or alanine showed a dramatic loss of affinity. (Muraki 2002)

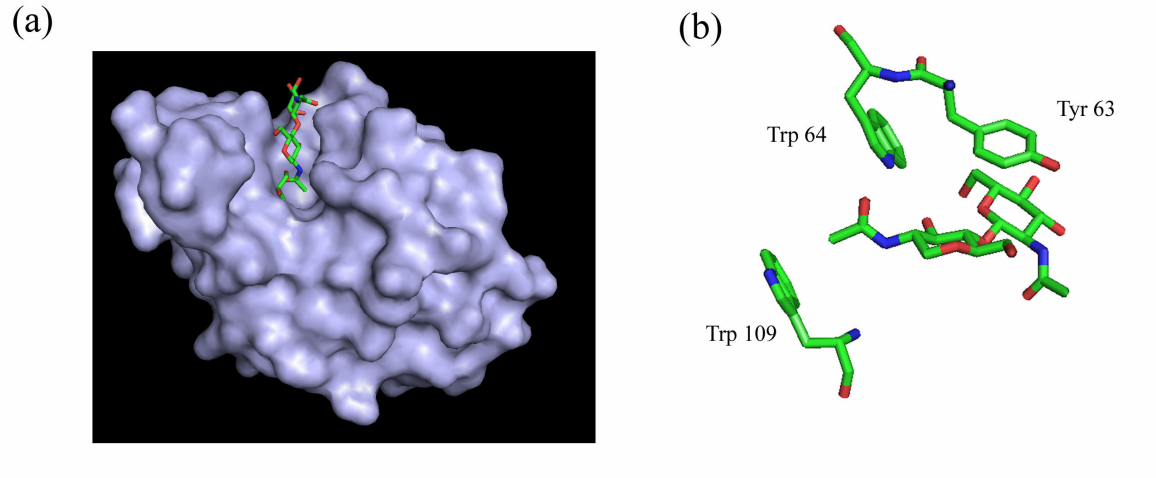


Figure 1.5 Human lysozyme with covalently bound GlcNAc(β1,4)GlcNAc. (a) Surface of lysozyme with GlcNAc(β1,4)GlcNAc in binding pocket. (b) Aromatic residues in binding pocket interacting with GlcNAc(β1,4)GlcNAc. (Muraki 1996, PDB: 1REY)

While aromatic-aromatic interactions are recognized as being important in protein-protein and protein-ligand interactions, their importance in pharmaceutical development has been vastly underappreciated. A large number of drugs contain aromatic substitutents that not only fit in hydrophobic pockets, but participate in aromatic-aromatic interactions that contribute to binding affinity. A clear example of this was found with the drug E2020, also known as Aricept, which was developed to treat symptoms of

Alzheimer's disease by binding to the enzyme acetylcholinesterase. The aromatic interactions within the active site with Aricept are shown in Figure 1.6. The crystal structure shows two parallel-displaced aromatic-aromatic interactions, a cation-π interaction, and an OH-π interaction with a bound water molecule. (Kryger 1998, Kryger 1999)

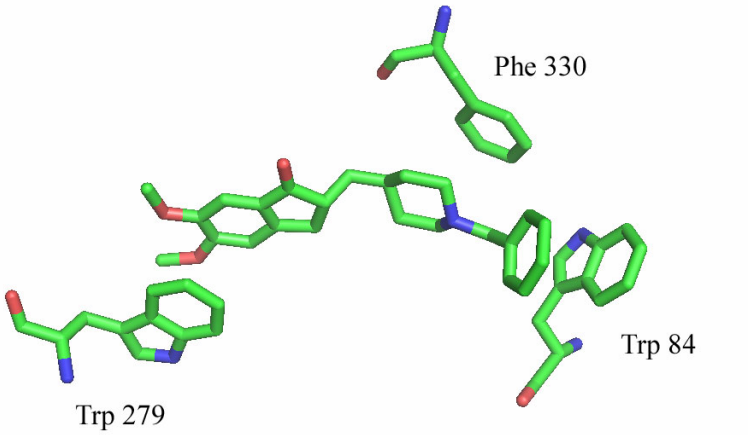


Figure 1.6 Aricept® in the binding pocket of acetylcholinesterase displaying aromatic-aromatic interactions with Trp 279, Phe 330, and Trp 84. (Kryger 1999, PDB: 1EVE)

A current area of drug development that could be greatly influenced by aromatic-aromatic interactions is the treatment of Alzheimer's disease and Type 2 diabetes. The development of both of these diseases is associated with the formation of amyloid fibrils. The fibrils associated with Type 2 diabetes are composed of islet amyloid polypeptide (IAPP). The shortest function fragment of IAPP is the sequence NFGAIL. It has been proposed that aromatic-aromatic interactions are important for the formation of fibrils in many amyloid proteins. (Gazit 2002) In IAPP, mutagenesis studies of the sequence gave similar or higher rates of aggregate formation for all the substitutions except the Phe to Ala mutant, which did not form any detectable aggregates. The authors propose the

development of drugs that block the π interaction to inhibit formation of amyloid fibers. (Azriel 2001) Several small molecules, shown in Figure 1.7, have been designed to control amyloid formation of the Aβ fibrils associated with Alzheimer's disease. In addition, the green birefringence of congo red dye (Figure 1.7c) has been used to characterize amyloid fibers. Interestingly, congo red has the same binding site as the small molecule Ro 47-1816/001 and is also a generic amlyoid inhibitor. (Lorenzo 1994)

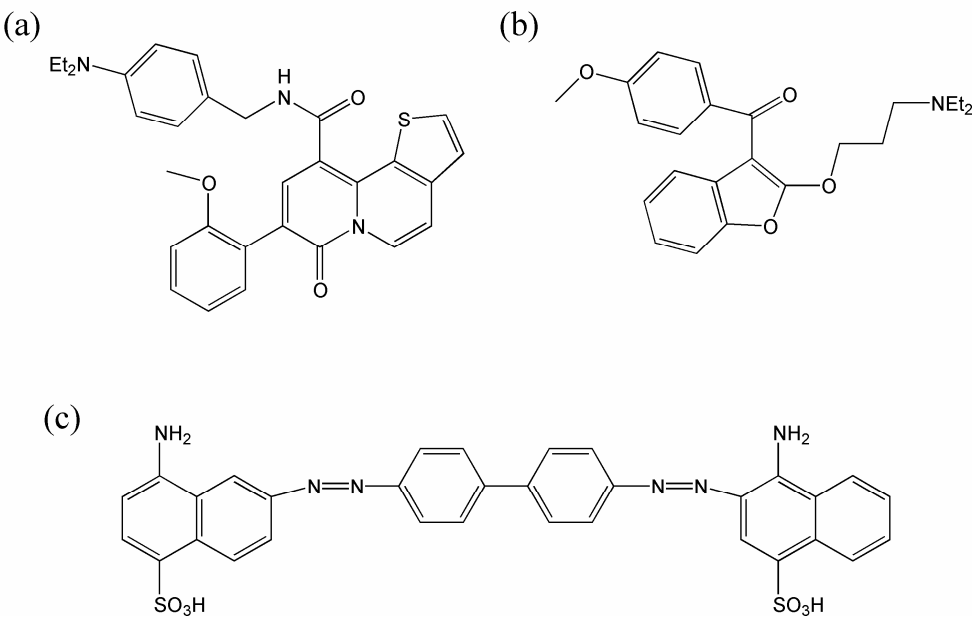


Figure 1.7 Structure of amyloid inhibitors. (a) Ro-47-1816/001 (Kuner 2000) (b) 3-p-toluoyl-2-(4'-(3-diethylaminopropoxy)-phenyl)-benzofuran (Twyman 1999) (c) congo red dye

The most striking example of aromatic-aromatic interactions in biological systems is the abundance of aromatic residues at the binding site of antibodies. At the interface of the A6 antibody and its antigen, the interferon γ receptor, there are 13 aromatic residues: six Tyr, six Trp, and one His. A total of ten of the residues out of 13 are in the binding pocket of the antibody; four in the V_L chain and six in the V_H chain. In a mutagenesis study of two of the three aromatic residues in the interferon γ receptor, replacement of a

tyrosine by alanine resulted in a reduction in binding of two orders of magnitude. Similarly, replacement of a tryptophan by an alanine resulted in an equally large reduction. Interestingly, mutation of the Tyr or Trp residues to Phe only resulted in a 6-fold reduction. (Hofstädter 1999)

1.2 AROMATIC-AROMATIC INTERACTIONS IN ABIOTIC FOLDING

The design of abiotic systems that utilize aromatic-aromatic interactions has increased rapidly since Hunter and Sanders published simplified rules to predict the favored geometry of aromatic interactions. (Hunter 1990, Hunter 2001) The six rules provided were based on the observation that aromatic-aromatic interactions occur when attractive interactions between π-electrons and the σ-framework outweigh unfavorable π-electron repulsion. While these rules were based on observations of porphyrin-porphyrin interactions, there were used to predict the experiment outcome of many abiotic systems that had been published previously.

Meta-phenylene ethynylene oligomers designed by Moore and co-workers form stable helical structures dictated by solvophobic-driven aromatic-aromatic interactions. The folding can be driven by solvent and temperature with the helical structure dictated by backbone constraints as shown in Figure 1.8. (Nelson 1997, Lahiri 2000) A variety of spectroscopic studies showed the oligomer folds into a helical conformation in more polar solvents such as acetonitrile and less polar solvents such as chloroform acts as a denaturant. The helical oligomers have been used to bind a rod-like diphenylpiperazine guest in the internal cavity. (Tanatani 2001) In addition, use of chiral side chains in the oligomer in a mixture with achiral oligomer showed transfer of chirality through cooperative intermolecular self-assembly of the helices into helical columns. (Brunsveld 2001) A water soluble oligomer was synthesized to take full advantage of the solvophobic driving force in polar solvents. Interestingly, increasing proportions of water

in acetonitrile increased the overall folding of the oligomer in up to 90% water-10% acetonitrile, with an unexpected deviation from the trend at 100% water. It was proposed that acetonitrile may be required to solvate the helical conformation and complete absence may cause an alteration in conformation that minimizes unfavorable interactions in water. (Stone 2004)

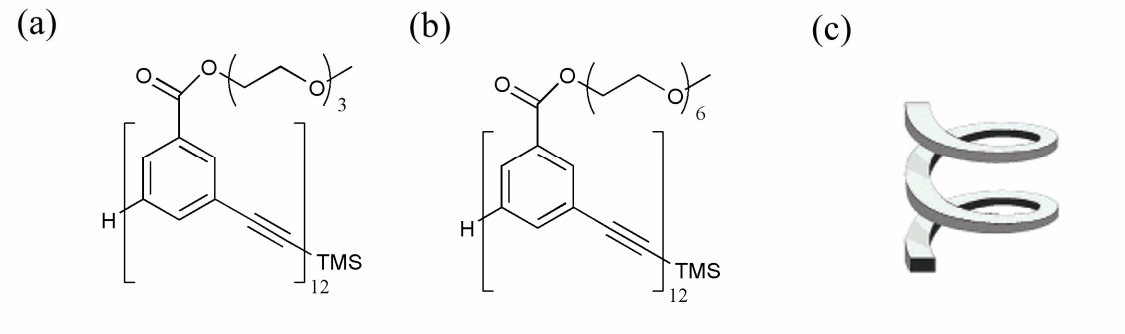


Figure 1.8 Moore's *meta*-phenylene ethynylene oligomers. (a) Original organic soluble oligomer (b) Water soluble foldamer (c) Cartoon of the helix fold of the oligomers. (Stone 2004)

Aromatic-aromatic interactions have been used to afford folding in a variety of other abiotic molecules. Pyrimidine-hydrazone oligomers have been used to promote helical structures using aromatic-aromatic interactions with imine linkages (Figure 1.9). (Gardinier 2000) Peptoid oligomers with chiral aromatic side chains were shown to promote formation of helices very similar to those of natural peptides in acetonitrile but without the hydrogen bonding that normally stabilizes them as shown in Figure 1.9b. The most stable helices resulted from creation of an "aromatic face" on the helix that maximized the aromatic-aromatic interactions without overcrowding from additional aromatic units. (Wu 2001) A columnar, folded trimer was created using aromatic-aromatic interactions between tethered, crowded hexa-substituted benzene units as shown in Figure 1.9c. NMR analysis confirmed the expected folded conformation in chloroform. (Zhang 2003)

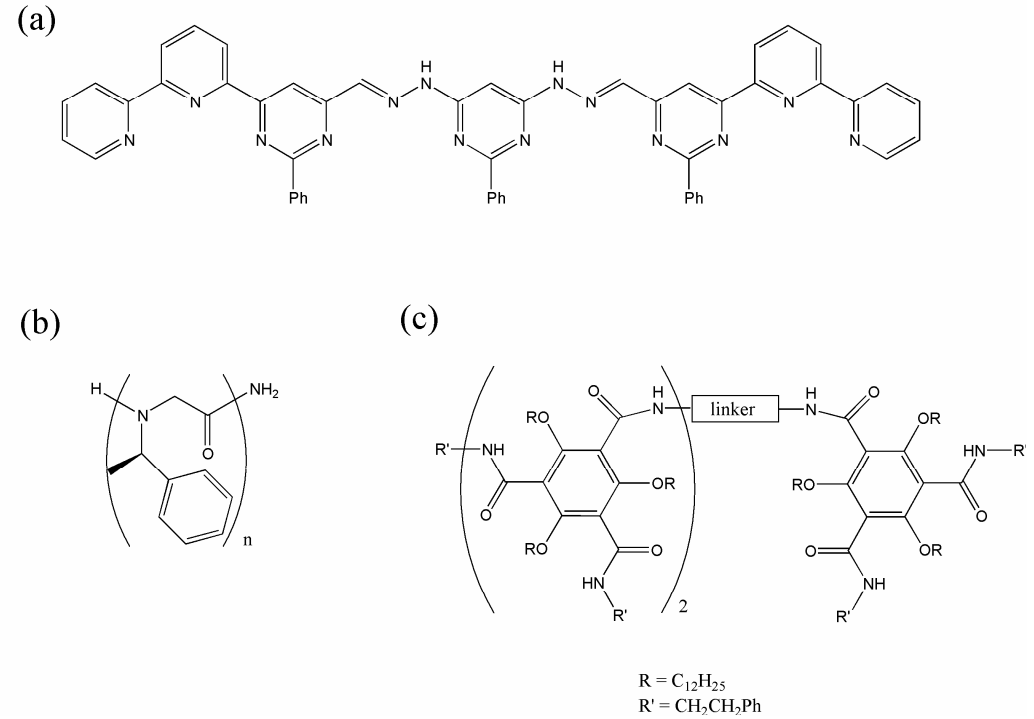


Figure 1.9 Examples of aromatic-aromatic interactions in abiotic folding. (a) Pyrimidine-hydrazone oligomers with imine linkages that promote helical structures. (Gardinier 2000) (b) Peptoid oligomers with chiral aromatic side chains. (Wu 2001) (c) Tethered, crowded benzene units that create a columnar face-centered fold. (Zhang 2003)

1.3 DONOR-ACCEPTOR TYPE AROMATIC-AROMATIC INTERACTIONS

1.3.1 Aromatic Donor-Acceptor Geometry

While most of the aromatic-aromatic interactions discussed thus far have utilized one of the four geometries shown in Figure 1.1 (parallel displaced, parallel staggered, T-shaped, and herringbone), aromatic-aromatic interactions of the donor-acceptor type are different. The electron-rich (donor) aromatic unit has increased electron density in the π-system and, therefore, a more negative partial charge as shown in Figure 1.10. The electron-deficient (acceptor) aromatic unit has a decreased density in the π-system and a more positive partial charge. When used together, they produce an aromatic-aromatic

interaction that has electrostatic complementarity which prefers a face-to-face geometry as shown in Figure 1.10c. This face-to-face geometry has the ability to shield a larger proportion of the hydrophobic aromatic face from water, thereby increasing the magnitude of the interaction as compared with other geometries.

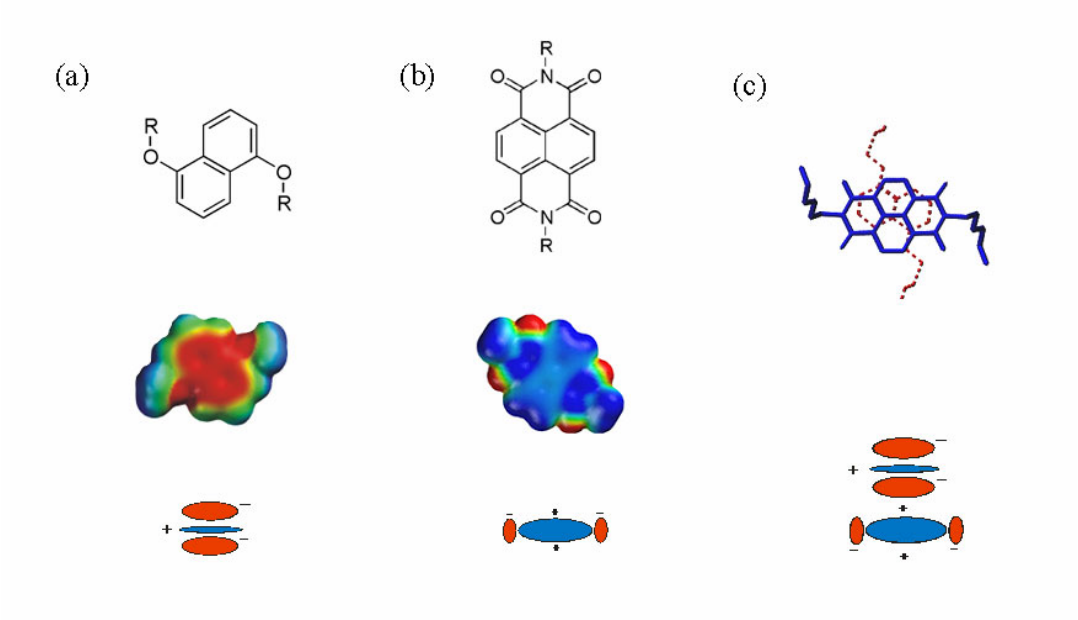


Figure 1.10 Donor-acceptor type aromatic-aromatic interactions. (a) Structure of the electron-rich aromatic unit 1,5-dialkoxynapthalene (DAN) with a visualization of the surface potential and cartoon model. (b) Structure of the electron-deficient 1,4,5,8-naphthalenecarboxylic diimide (NDI). (c) Crystal structure of the DAN-NDI complex and cartoon of the electrostatic complementarity. (Lokey 1995)

1.3.2 Donor-Acceptor Type Aromatic-Aromatic Interactions in Biology

While donor-acceptor type aromatic-aromatic interactions have been used in abiotic systems for many years, the first implication of this type of aromatic-aromatic interaction in biology has been often overlooked. As mentioned earlier, aromatic-aromatic interactions contribute substantially to the antibody-antigen interaction as a result of the large number of aromatic residues, especially tyrosine, on the surface of the binding region of antibodies. However, the aromatic amino acids are generally electron

rich. During early research into the nature and specificity of antibodies in the 1930's, Landsteiner used a variety of benzene derivatives, the majority of which were electron-deficient to sensitize animals. The anaphylactic response to these compounds was gauged to determine the specificity of the response among structurally similar compounds. All the multiple nitrated benzene derivatives elicited very strong responses to skin sensitization, almost assuredly from the donor-acceptor aromatic-aromatic interactions with the electron-rich aromatic residues in the binding pocket of the antibodies. (Landsteiner 1935)

1.3.3 Aromatic Donor-Acceptor Interaction in Rotaxanes and Catenanes

Donor-acceptor aromatic-aromatic interactions have been used by Stoddart and co-workers in catenanes and rotaxanes for a variety of purposes as shown in Figure 1.11. Much of the work began with a tetracationic cyclobis(paraquat-p-phenylene) cyclophane host in combination with a variety of electron-donating aromatic guests. The donor-acceptor aromatic-aromatic interactions have been shown to be useful in the templated synthesis of rotaxanes via either clipping, or threading-and-stoppering. (Moonen 2005) Donor-acceptor aromatic-aromatic interactions had been used in the self-assembly of [2]- and [3]-pseudorotaxanes selectively. (Ashton 1991) Bistable rotaxanes were used to shuttle between two different aromatic donor-acceptor interactions by either chemical or electrochemical switching methods with an electron-rich cyclic dialkoxynaphthalene (DAN) dimer. (Vignon 2004, Iijima 2004) The cyclic DAN had a preference for the naphthalene diimide (NDI) unit of the rotaxane. However, the linker for the cyclic DAN dimer is composed of 12-crown-4 linkers that have a known affinity for lithium ions. Addition of lithium ions was shown to switch the binding site to the smaller pyromellitic diimide (PDI) unit in order to accommodate binding of the ions to the cyclic DAN linkers. (Vignon 2004) In addition, cyclic voltammetry experiments showed that the

switching could be performed electrochemically, as binding affinity to the NDI unit is diminished upon one-electron reduction. (Iijima 2004)

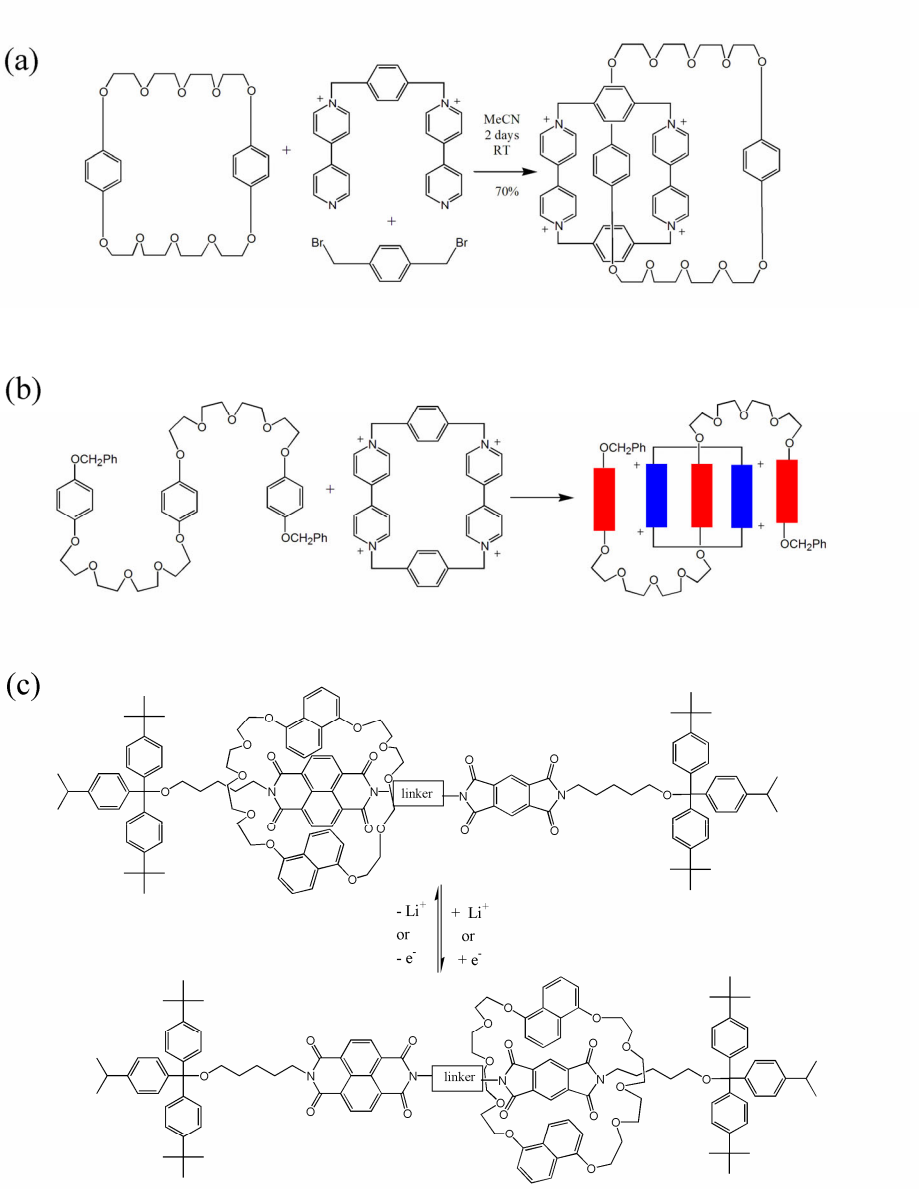


Figure 1.11 Donor-acceptor aromatic-aromatic interactions in rotaxanes and catenanes. (a) Templated synthesis of a catenane. (Claessens 1997) (b) Self-assembly of a pseudorotaxane. (Ashton 1991) (c) Switching of a bistable rotaxane by lithium ions or reduction/oxidation. (Iijima 2004)

Switching of bistable rotaxanes has also been accomplished by changes in pH and a light-fueled oxidation of an electron-rich tetrathifulvalene unit, although the latter was complicated by substantial folding and undesired aromatic-aromatic interactions. (Moonen 2005, Saha 2007) Bistable rotaxanes have been investigated for use in molecular machines as molecular muscles and nano-valves. (Northrop 2007) Donor-acceptor aromatic-aromatic interactions have also been used for template-directed synthesis of both charged and neutral catananes in addition to a shuttling process similar to that of rotaxanes. (Hamilton 1997, Amabilino 1995)

1.3.4 Aedamers

1.3.4.1 Donor-Acceptor Aromatic-Aromatic Interactions in Folding

*A*romatic *e*lectron *d*onor-*a*cceptor oligo*mers* (*aedamers*) developed in the Iverson group were the first folding molecules to use donor-acceptor aromatic-aromatic interactions to determine the secondary structure of a complex synthetic molecule in aqueous solution. (Lokey 1995) The molecule was designed to fold into a pleated structure in water by alternating electron-rich and electron-deficient aromatic units driven by the hydrophobic effect. The electron-rich 1,5-dialkoxynaphthalene (DAN) unit was linked alternately with the electron-deficient 1,4,5,8-naphthalenetetracarboxylic diimide (NDI) with aspartic acid residues as shown in Figure 1.12.

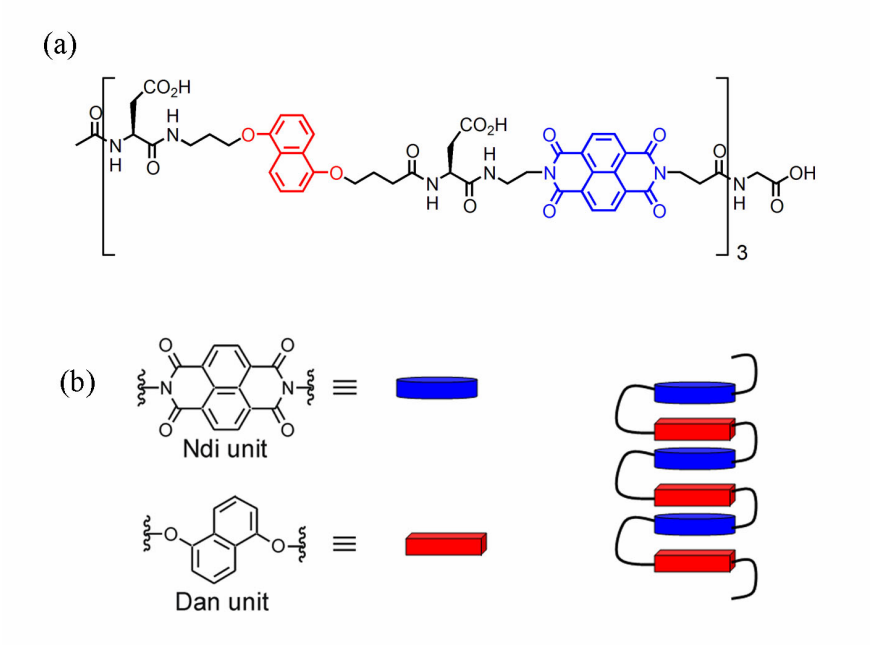


Figure 1.12 (a) Chemical structure of the aedamer. (b) Cartoon representation of the aedamer.

The pleated structure was verified by UV and visible absorption, as well as NMR. The face-to-face stacked aromatic core of the pleated structure allows for orbital mixing of the aromatic units, which is observed as a charge transfer band in the visible region. The charge transfer band gives the aedamer its characteristic merlot color. In addition, the DAN and NDI absorbance in the UV exhibits significant hypochromism compared to the independent monomers, similar to the hypochromism effect that is characteristic of DNA in a double helix. The hypochromism is dependent on the distance, r, as a function of $1/r^3$ and is highly orientation-dependent. Both of these spectroscopic signatures provide evidence for the face-to-face stacking of the oligomers. The pleated structure was also very robust. The charge transfer absorbance was relatively temperature independent, indicating the structure could not be unfolded at elevated temperatures in water.

The modular nature of the aedamer allowed for the design of an amphiphilic version, where the aspartate linkers on one side of the aedamer were replaced with leucine as shown in Figure 1.13a. (Nguyen 1999) This provided an amphiphilic pleated structure conceptually analogous the leucine zipper motif in proteins. As expected, the amphiphilic aedamer was highly aggregated in its folded form, unlike the original aedamer. It was discovered that upon heating to 80° C, the charge transfer absorbance disappeared and the aedamer became so highly aggregated that a hydrogel was formed as shown in Figure 1.13b. In addition, it was determined that the hydrogel formation was product-promoted. The model of aggregation proposed was one in which unfolding of the pleated, stacked structure was followed by random tangled aggregation, similar to that of collagen networks.

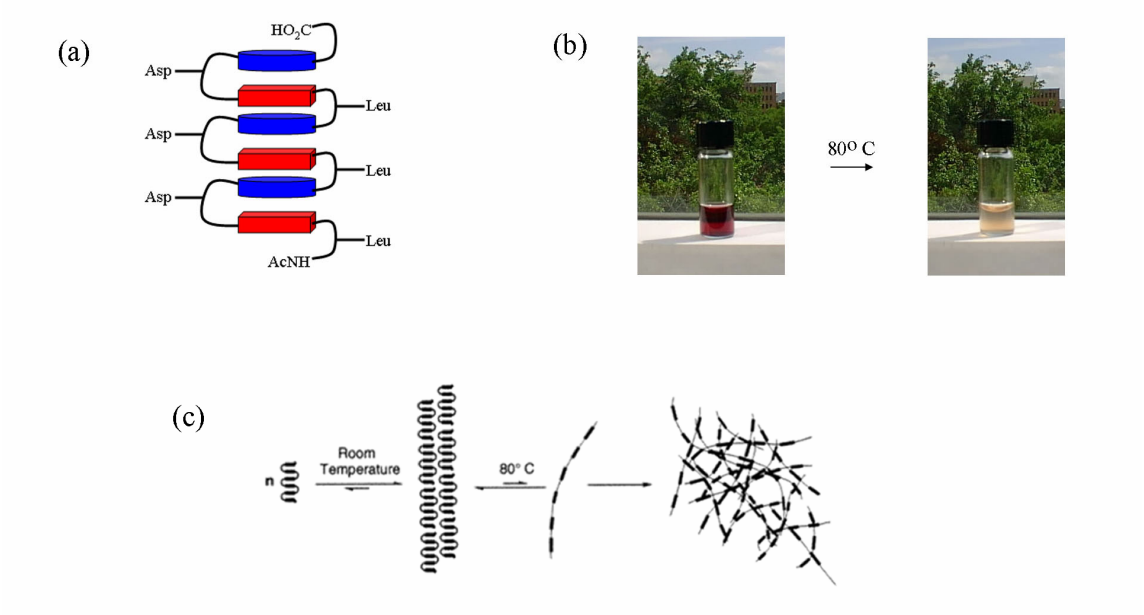


Figure 1.13 (a) Cartoon of the amphiphilic aedamer. (b) Loss of the charge transfer absorbance upon heating. (c) Proposed model of aggregation and hydrogel formation. (Nguyen 1999)

The structure of the pleated aedamer was studied in more detail by synthesizing a model dimer system with various linkers by NMR. (Zych 2000) The DAN and NDI

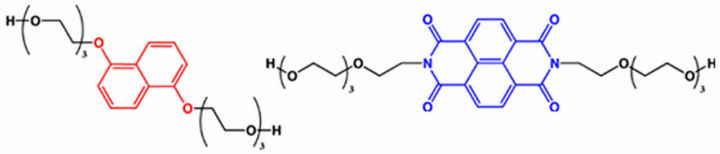
aromatic protons showed a diagnostic upfield shift, which was consistent with face-to-face stacking and ring current effects. In addition, NOE cross-peaks were observed between the aromatic protons in the DAN and NDI dimers. To analyze the orientation and stacked structure of the aromatic units, an algorithm was developed which simulated the chemical shift changes resulting from the distance and angle between the aromatic proton of interest and the shielding species, in computer modeled structures. The result showed a good fit between the actual NMR chemical shifts and the computer-modeled low energy structures. The computer modeling, however, gave an ensemble of nearly isoenergetic conformational states with stacked aromatic units. The results imply that the pleated structure has folded conformations that are dynamic in nature. Interestingly, a control molecule designed to prevent stacking of the DAN and NDI unit in the dimer was found to aggregate extensively in solution. This observation supported the model of aggregation for the amphiphilic aedamer in that aggregation results from unfolded conformations.

The detailed structure of the aedamer was explored further by probing the conformational modularity of dimer units. (Zych 2002) The study was designed to determine the persistence of conformational propensities of the dimer units in larger trimer and tetramer structures. While not all the systems tested displayed conformational modularity, a DAN-NDI-DAN trimer composed of two different linkers gave chemical shifts similar to the prediction from the DAN-NDI and NDI-DAN dimers individually. The systems that did not match predictions well were explained in computer modeling by a steric clash of one of the dimer linkers with the additional DAN unit in the trimer.

In addition to the exploration of the detailed structure of the interaction, a more fundamental understanding of the driving force and solvent dependence of the NDI-DAN interaction was required. This was accomplished by a detailed investigation of the

association of DAN and NDI monomers in solvents with a variety of polarities by NMR as shown in Figure 1.14. (Cubberly 2001) Monomers with tetraethylene glycol side chains for both the DAN and NDI were used. The interaction could have been driven strictly by the hydrophobic effect, electrostatic complementarity, or both. The results showed that the stability of the complex increased with increasing solvent polarity with a nearly linear correlation between the association and solvent polarity. The association of NDI-DAN monomers showed an increase of nearly two orders of magnitude between methanol and water.

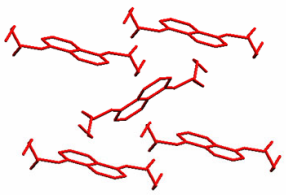
(a)



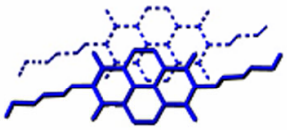
(b)

Solvent	K_a (M^{-1}) Dan:Dan	K_a (M^{-1}) Ndi:Ndi	K_a (M^{-1}) Dan:Ndi
$CDCl_3$	-	-	2
DMSO-*d6*	1	2	3
CD_3CN	1	3	11
CD_3OD	1	8	30
D_2O	20	245	2045

(c)



(d)



(e)

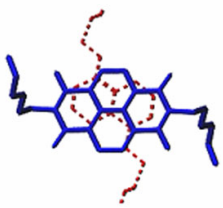


Figure 1.14 (a) NDI and DAN monomers used to study the effect of solvent on the NDI-DAN interaction. (b) The association constants of the self-association of the DAN and NDI monomers and the NDI-DAN complex in various solvents. (c) Crystal structure of DAN monomer. (d) Crystal structure of NDI monomer. (e) Co-crystal of DAN-NDI complex. (Cubberly 2001)

The self-association of the DAN monomers was two orders of magnitude less and the self-association of the NDI was one order of magnitude less than the NDI-DAN association. This was explained to be primarily due to stacking geometry. The NDI-DAN association has a face-to-face geometry which maximizes the desolvation of the hydrophobic surface area (Figure 1.14e). The NDI self-association is moderate because the geometry is a parallel displaced, stacked conformation which desolvates a large portion of the aromatic faces (Figure 1.14d). The DAN self-association is small (20 M-1) due to the edge-face conformation which limits the desolvation (Figure 1.14c). The results support the hypothesis that the driving force of NDI-DAN interaction is a classic hydrophobic effect.

With a greater understanding of the aedamer structure and driving force, the ability of the modular nature of the system to control structure was probed. In a manner analogous to the site-directed mutagenesis of proteins, the primary structure of the aedamer was altered to determine whether secondary structure could be affected to afford new folding topologies. The pleated fold was changed to an intercalative fold, similar to a turn structure in proteins, by using non-alternating NDI and DAN units as shown in Figure 1.15. (Gabriel 2005) A trimer consisting of the pleated structure, DAN-NDI-DAN, was compared to the intercalative structure, DAN-DAN-NDI, in addition to a control trimer with a shorted linker between the two DAN units so that it only has the ability to fold into a dimer. The pleated and intercalative folded trimers exhibited similar charge transfer absorbance and hypochromism, whereas the control trimer gave values similar to a dimer indicating its inability to fold completely. This data supported the face-to-face stacked orientation of the intercalative folded trimer.

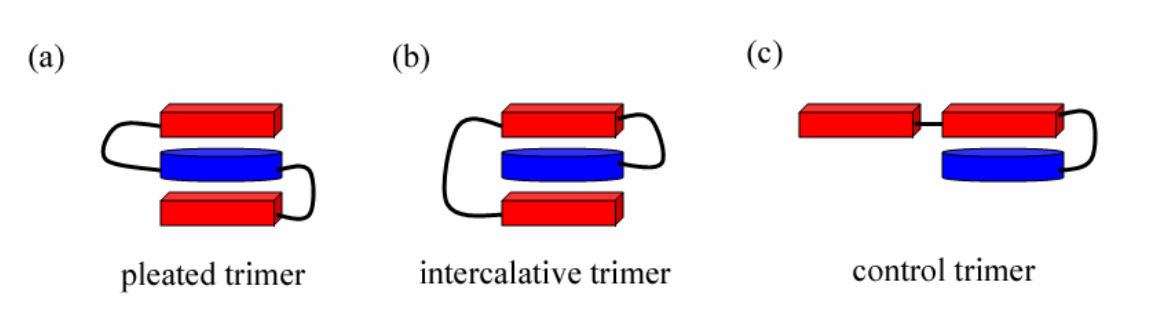


Figure 1.15 Cartoons of trimers used to study the ability of the NDI-DAN interaction to control secondary structure. (a) The pleated, alternating trimer. (b) The intercalative, non-alternating trimer. (c) The control trimer that is prohibited from folding by a shorted linker.

NMR showed similar upfield shifting of all the aromatic protons of the intercalative folded trimer compared to the pleated trimer. In addition, NOESY experiments gave NOE cross-peaks between the central NDI unit and both DAN units of the intercalative folded trimer. Computer modeling using NOE constraints were consistent with the stacked structure of the aromatic units in the intercalative fold. The modeling suggested dynamic conformations of the trimer very similar to what had been discovered for NDI-DAN dimers. (Zych 2001)

Several other synthetic systems whose folding is controlled by donor-acceptor aromatic-aromatic interaction have been developed in recent years. A δ-peptide foldamer based on ornithine was developed. The side chains on the α-nitrogen of ornithine was modified with PDI and DAN side chains as shown in Figure 1.16. The peptide was then synthesized with a trimer of PDI followed by a trimer of DAN. The donor-acceptor aromatic-aromatic interaction creates a turn between the blocks of PDI and DAN to create a "zipper" foldamer in organic solvents. The folded conformation exihibited characteristic hypochromism, NMR upfield shifts, and a significant circular dichroism signal. (Zhao 2004)

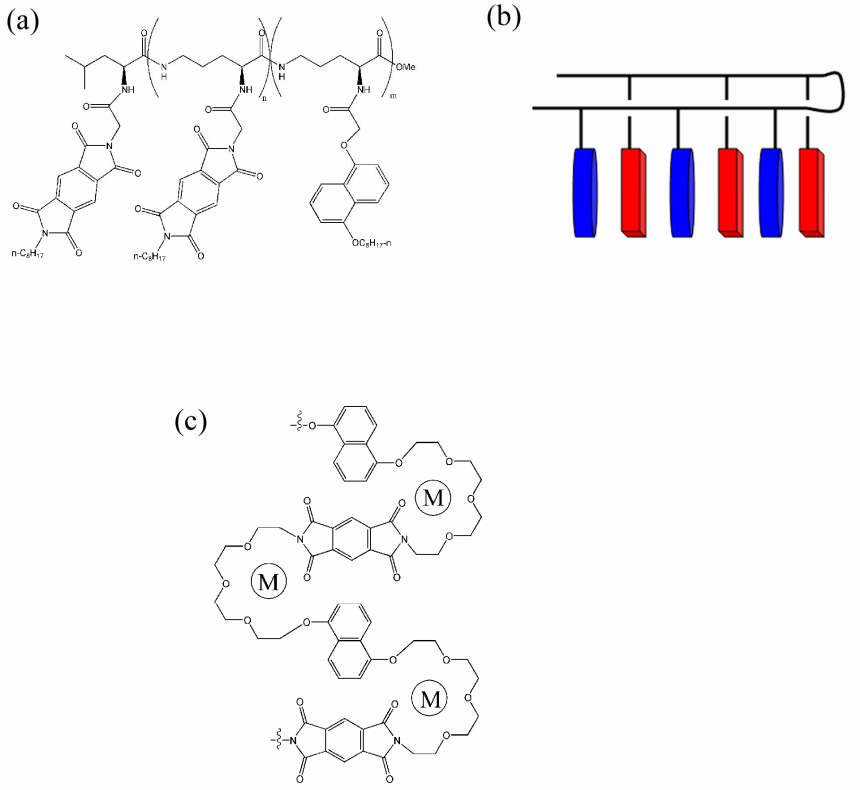


Figure 1.16 Other synthetic systems whose folding is controlled by donor-acceptor aromatic-aromatic interactions. (a) A δ-peptide foldamer that forms a turn to maximize the donor-acceptor aromatic-aromatic interaction. (b) A cartoon of the resulting zipper structure. (Zhao 2004) (c) An alternating polymer whose folding is governed by donor-acceptor aromatic-aromatic interactions. (Ghosh 2004)

A polymer synthesized of alternating DAN and PDI units was synthesized using polyethylene oxide linkers of various lengths as shown in Figure 1.16c. The polymer exhibited classic solvophobic behavior as shown by the charge transfer absorbance and the NMR upfield shifts, although the polymer was not water-soluble. The polymer was significantly folded in solution, however the addition of alkali metal ions increased the degree of folding. The polymer synthesized with a tetraethylene glycol linker had the highest degree of folding without the addition of alkali salts, whereas the polymer with a

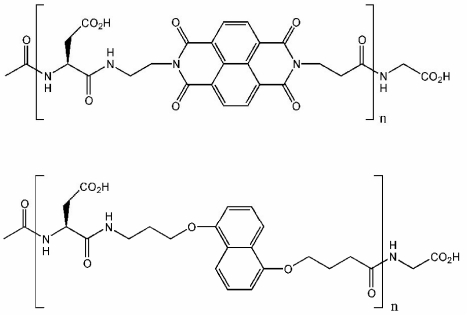
pentaethylene glycol linker folded with the addition of lithium salts and hexaethylene glycol linked polymers folded with the addition of potassium salts. (Ghosh 2004, Ghosh 2005a) A polymer of PDI folded into a similar structure to the alternating polymer with the addition of DAN monomers with cationic amine side chains. (Ghosh 2005b)

1.3.4.2 Donor-Acceptor Aromatic-Aromatic Interactions in Intermolecular Association

A natural extension of the *intra*molecular donor-acceptor aromatic-aromatic interactions in the aedamer system is the exploration of using donor-acceptor aromatic-aromatic interactions to direct self-assembly in an *inter*molecular system. Oligomers up to four units of only DAN units and only NDI units were designed to direct the assembly of a heteroduplex in water as shown in Figure 1.17. (Gabriel 2002) The interaction of the independent oligomers was probed by isothermal titration calorimetry (ITC) and NMR titrations. The formation of duplexes was energetically favorable and showed an increase in association of an order of magnitude per additional association from monomers to tetramers.

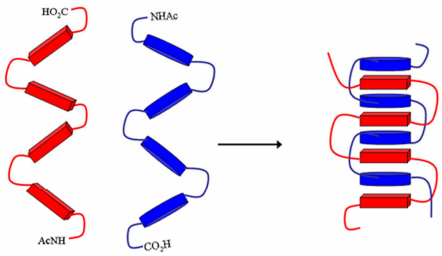
The association constant of tetramer strands was 3.5×10^5 M^{-1} as compared to 130 M^{-1} for the monomers. The association constant of the tetramers is significant considering the large charge repulsion between the strands. Each aromatic unit in a strand uses a negatively charged aspartate linker and the *C*-terminus of the strand is also a negatively charged carboxylate. In the tetramer, the association brings ten negative charges in close proximity, yet the association constant is still considerably large.

(a)



n = 1- 4

(b)



(c)

Duplex	K_a (M^{-1})	ΔG° (kcal/mol)
n = 1	1.3×10^2	-3.1
n = 2	2.7×10^3	-5.0
n = 3	4.5×10^4	-6.8
n= 4	3.5×10^5	-8.1

(d)

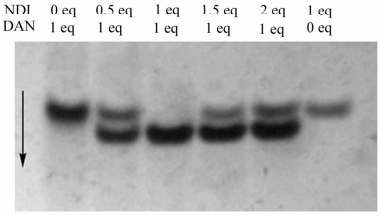


Figure 1.17 (a) The chemical structures of the independent NDI and DAN strands of the heteroduplex. (b) A cartoon of the heteroduplex formation. (c) The association constants and energies associated with heteroduplex formation. (d) PAGE titration indicating specificity of the heteroduplex formation (Gabriel 2002)

In order to probe the heteroduplex formation and the specificity of interaction, a titration experiment was performed on a polyacrylamide gel as shown in Figure 1.17d. At a 1:1 ratio, a single band is seen with no streaking, indicating a single discrete complex has been formed. In addition, excess oligomer runs slower down the gel indicating no complexes larger than a heteroduplex has been formed. The success of the PAGE experiment was a testament to the strength of the heteroduplex association as weakly associating species do not run as the intact complex on a polyacrylamide gel.

Several other synthetic systems have been designed to utilize donor-acceptor aromatic-aromatic interactions in an intermolecular fashion. A similar artificial duplex was designed using PDI and DAN strands, albeit in organic solvents as shown in Figure

1.18. The binding constant of the tetramer strands displayed a binding constant of 8.8 x 10^3 M^{-1}. (Zhou 2003) In an attempt to improve the binding constant in more polar organic solvents, a more polar linkage was used. Interestingly, the binding constant of the tetramer strands did not improve (4.7 x 10^3 M^{-1}). However, the use of a tetramer of PDI with a pentamer of DAN doubled the Ka to 1.02 x 10^4 M^{-1}. (Zhou 2005)

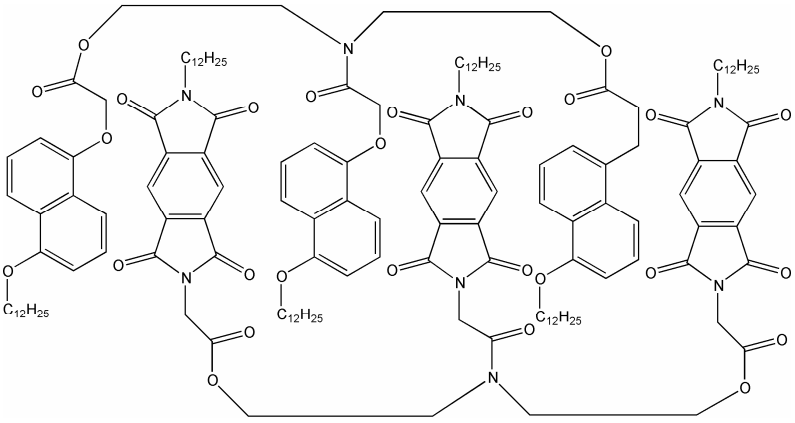


Figure 1.18 Artificial duplex in organic solvents. (Zhou 2003)

An interesting use of donor-acceptor aromatic-aromatic interaction in an intermolecular association was the design of a β-hairpin with tryptophan residues in positions diagonally across the strands as shown in Figure 1.19. (Butterfield 2004) The hairpin was shown to bind flavin mononucleotide (FMN), presumably in the face-to-face conformation as predicted by computer modeling. Interestingly, the reduced form of the flavin mononucleotide showed an order of magnitude reduction in association compared to the oxidized form (40 M^{-1} versus 720 M^{-1}). In addition, the flavin interaction was greater than the interaction of ATP with the same hairpin. This difference was attributed to the increased electron-deficient nature of the oxidized flavin compared to both the reduced form of FMN and adenine in ATP, therefore increasing the favorable electrostatic complementarity with the electron-rich tryptophan.

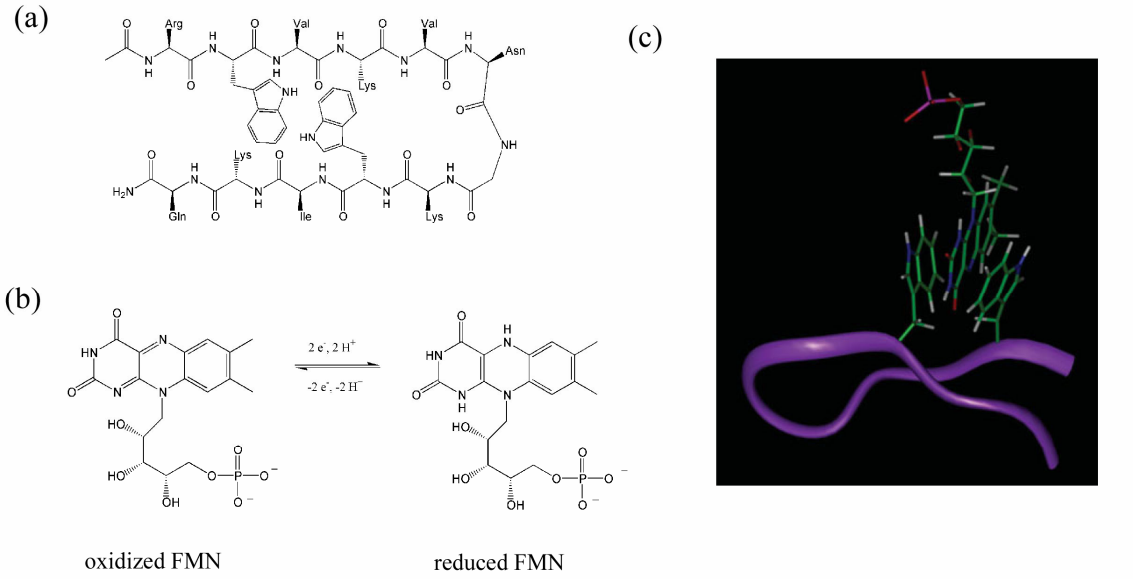


Figure 1.19 (a) β-Hairpin designed to bind flavin mononucleotide (FMN) through aromatic-aromatic interactions with diagonal, cross-strand tryptophan residues. (b) Structure of the oxidized and reduced form of FMN. (c) Computer model displaying the face-centered stacking of the Trp-FMN-Trp interaction. (Butterfield 2004, Waters 2004)

1.3.4.3 Donor-Acceptor Aromatic-Aromatic Interactions in Materials

While the intermolecular association of independent strands of NDI and DAN oligomers was relatively high affinity in water, it was proposed that incorporating NDI and DAN units into independent water-soluble polymer strands would produce extremely high affinity intermolecular macrostructures. This high affinity could give rise to interesting material properties. DAN and NDI polymers were synthesized by functionalization of polyethylene-*alt*-maleic anhydride (Figure 1.20). (Reczek 2006a) The polymers were soluble in basic solution as a result of the large number of negative charges on the polymer chains. Mixtures of the polymer strands showed different behavior than the independent strands, both in solution and in the solid state. In solution, interaction of the DAN and NDI units in the polymer mixtures was confirmed by the

presence of a charge transfer absorbance. The polymer mixture also displayed a 5-fold increase in viscosity over either of the independent strands, signifying a large number of interchain interactions.

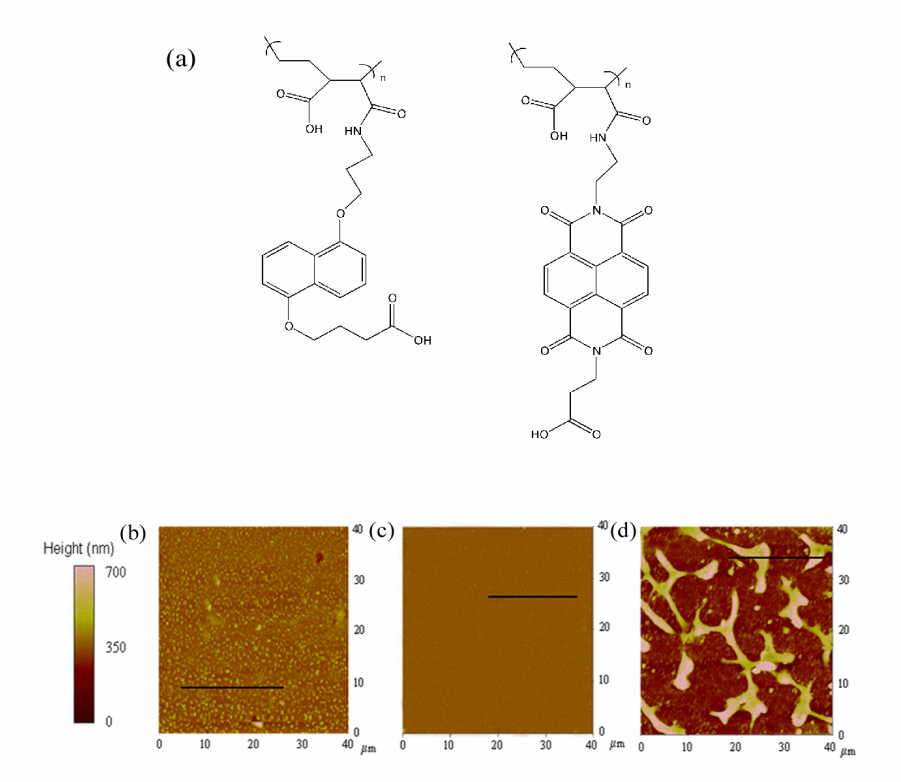


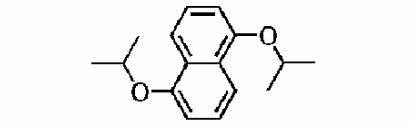
Figure 1.20 (a) Structure of NDI and DAN containing polymers. (b) AFM of a thin film of the DAN polymer. (c) AFM of a thin film of the NDI polymer. (c) AFM of a thin film of the mixture of DAN and NDI polymers. (Reczek 2006a)

The investigation of thin films of the independent strands compared to the mixture by AFM showed the NDI strands alone formed smooth thin films, and the DAN strands formed micelles, while the mixture displayed large macrostructures characteristic of an integrated polymer network, as shown in Figure 1.20b-d. Additionally, the mixture could be drawn into fibers that appeared to be made of densely packed threads.

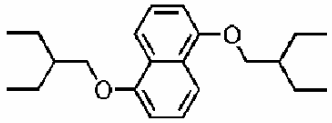
While previous studies had been performed on the DAN-NDI monomer interaction in solution, additional studies on the DAN-NDI interaction were performed in

the absence of solvent in order to study the bulk material properties of a mixture of monomers. (Reczek 2006b) A series of DAN and NDI monomers were synthesized with variations in the length, chirality, and steric bulk of the side chains (Figure 1.21). The DAN-NDI mixtures produced mesophases that were stable over large temperature ranges from 25° C to 110° C. The charge transfer band was observed in the mesophases indicating assembly of extended columns with alternating face-to-face stacking of the NDI and DAN units. A correlation was discovered between the clearing point of the mixture and the melting/clearing point of the NDI component as shown in Figure 1.21b. In addition, the crystallization points of the mixtures coincided with the DAN melting points. These trends are consistent with the NDI side chains having more restricted motion than the DAN side chains in the mesophase so that the crystallization point coincides with the restriction of motion of the DAN side chains. X-ray powder diffraction showed that the unit cell dimensions of the mesophase are controlled by the length and size of the DAN and NDI side chains. Use of sterically hindered derivatives in mixtures showed a dramatic and sudden color change from a deep red to yellow upon crystallization of the mesophase due to a phase separation of the NDI and DAN units as shown in Figure 1.21c.

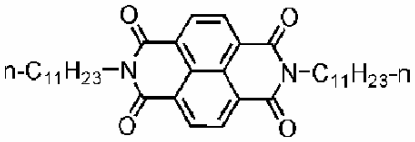
(a)



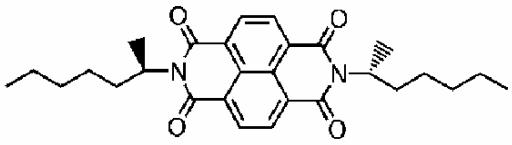
DAN-iPr_2



DAN-EBu_2



NDI-Und_2



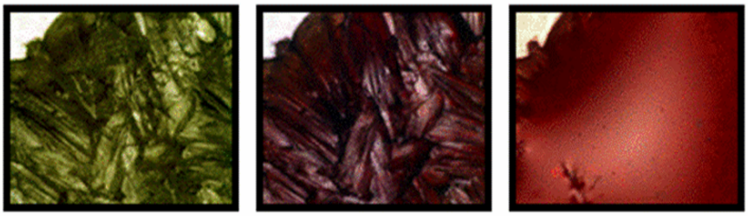
NDI-(R)-$MeHex_2$

(b)

DAN melting points	NDI clearing points: NDI-Und_2	NDI clearing points: NDI-(R)-$MeHex_2$
	162°	159°
DAN-iPr_2 101°	140° - 79°	129° - 81°
DAN-EBu_2 31°	136° - 26°	123° - 27°

(c)

DAN-iPr_2
+
NDI-(R)-$MeHex_2$



Crystalline Phase | Mesophase | Isotropic Phase

Figure 1.21 (a) Structures of NDI and DAN monomers used for mesophase formation. (b) Clearing (left) and crystallization (right) points of the mesophases of NDI-DAN mixtures in comparison to the clearing points of the NDI monomers and melting points of the DAN monomers used in the mixture. (c) Color of 1:1 mixtures of DAN-iPr_2 and NDI-R-$MeHex_2$ in the crystalline phase (60° C), mesophase (110° C), and isotropic phase (160° C). (Reczek 2006b)

1.4 Overview of Peptide-Based Donor-Acceptor Aromatic-Aromatic Interactions Projects

The work involving donor-acceptor aromatic-aromatic interactions which direct intermolecular self-assembly has allowed the design of many of the systems that will be described in this dissertation. The intermolecular association of the NDI and DAN oligomers in water lends itself to potential uses in biological systems. Chapter 2 describes the refined model of aggregation of an amphiphilic aedamer, which behaves similarly to naturally occurring amyloid proteins, known to be the pathological findings of many diseases. While it does not use the same non-covalent interactions to fold as natural proteins, this abiotic foldamer shares many of the same folding properties with natural proteins.

The selective interaction of the DAN and NDI is orthogonal to other non-covalent interactions in proteins and peptides. In addition, the intermolecular association work revealed that the single interaction of DAN and NDI monomers is worth -3.1 kcal/mol. This value is comparable to that of a hydrogen bond, and may allow the use of donor-acceptor aromatic-aromatic interactions to alternatively stabilize protein and peptide structure as discussed in Chapter 3. In addition, the relatively high affinity of the tetramers can be exploited for biological use in protein expression tags as discussed in Chapter 4.

Chapter 2

Amyloid-Behavior in Abiotic, Amphiphilic Foldamers

2.1 CHAPTER SUMMARY

Introduction

There is a growing appreciation for the importance of alternative protein conformations. Biological activity is not restricted to globular folded structures because the polymeric nature of proteins combined with the unique conformational properties of the polypeptide backbone can lead to structured aggregates that have important consequences *in vivo* and *in vitro*. Previously, we reported an abiotic amphiphilic foldamer that, upon heating, undergoes an irreversible conformational change to a highly aggregated state. (Nguyen 1999) In this chapter, this previous work is extended through the study of a series of structurally related amphiphilic foldamers.

Goals

It is the aim of this chapter to further refine the model of conformational switching and aggregation of the amphiphilic foldamer. More specifically, the goals of this chapter are to determine how side chain structure and hydrophobicity affect folding behavior, determine some of the material properties of the aggregated states, and better define the structure of the aggregated species.

Approach

Several amphiphilic foldamers with alternate hydrophobic side chains were investigated for their aggregation behavior, UV signatures, and circular dichroism

spectra. Further, the material properties were determined by rheology and the microscopic structure was analyzed by SEM.

Results

Prior to heating, all foldamers of the series exhibited spectral characteristics consistent with folding in the pleated, stacked geometry characteristic of this class of foldamer. Following heating at 80°C, three of the four molecules exhibited irreversible aggregation to produce hydrogels. The hydrogels were characterized by rheology measurements and circular dichroism spectra revealed that hydrogel formation was dependent on highly ordered intermolecular assembly, conceptually analogous to protein amyloid formation. Hydrogel formation had the effect of amplifying the subtle structural differences between molecules, as the three amphiphilic foldamer constitutional isomers that formed hydrogels upon heating displayed significant differences in hydrogel properties. Taking a global view, our results indicate that amyloid-like behavior is not unique to proteins, but may be a relatively general property of amphiphilic folding molecules in aqueous solution.

2.2 BACKGROUND

Hydrogels have been widely studied due to the variety of applications of their unique properties. Hydrogels are used commercially in personal care and food products in addition to drug delivery applications. (Park 1996, Burdick 2005, Park 1993, Haines 2007) Polymer hydrogels, either covalent or non-covalently networked, are often the first choice for these applications. Recently, biologically-derived hydrogels have gained interest. Most commonly, peptides have been utilized to achieve gelling behavior, although a DNA hydrogel has also been reported. (Um 2006) Low molecular weight

hydrogelators (LMWG) and peptide hydrogels generally fall into two categories of behavior: chemical switching and physical switching. Chemical switching arises from chemical modification of a structure to induce a gelling state, such as hydrolysis, enzymatic cleavage, or covalent cross-linking and is generally irreversible. (Toledano 2006, Yang 2007, Yang 2006, Zhang 2003) Physical switching involves altering the conditions such that non-covalent interactions are formed or altered and such switching is often reversible. (de Loos 2005, Shen 2007, Ramachandran 2006, Hart 2007, Schneider 2002, Collier 2001, Yan 2006, Deming 2005) Examples of different types of hydrogels are given below in Figure 2.1.

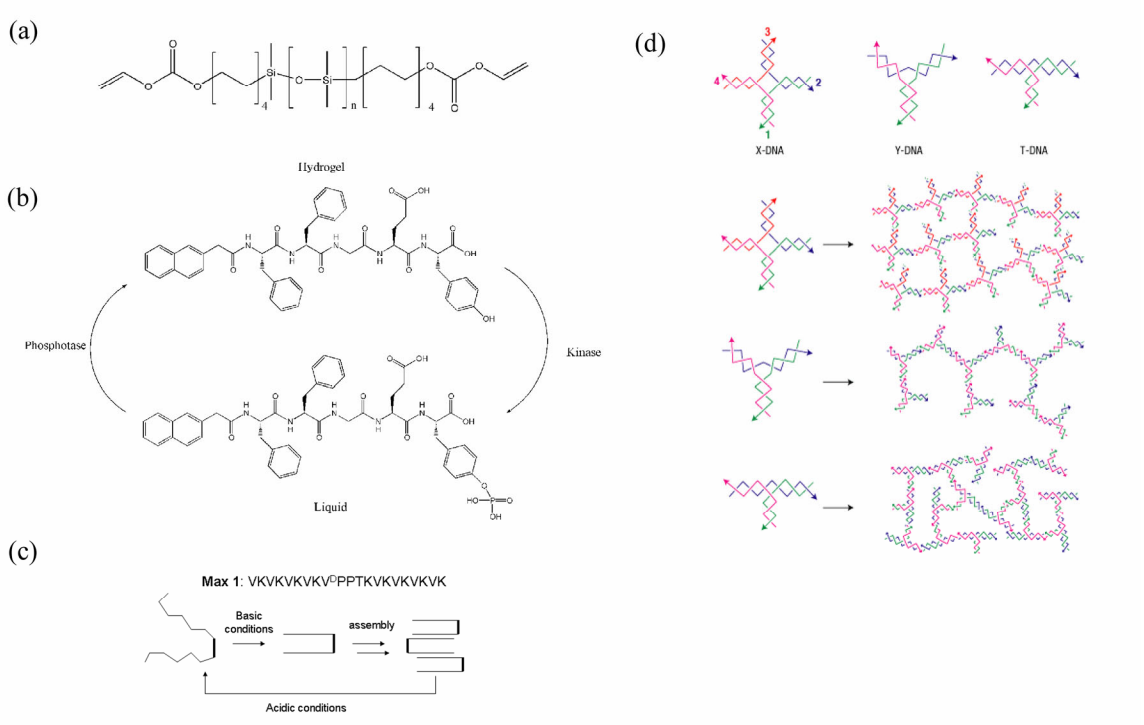


Figure 2.1 Types of hydrogels. (a) Polymer hydrogel in Balaficon A soft contact lenses (Bausch and Lomb). (b) Peptidic hydrogel with chemical switching by enzymatic phosphorylation/dephosphorylation. (Yang 2006) (c) Peptidic hydrogel with physical switching by pH. (Schneider 2002) (d) DNA hydrogel. (Um 2006)

There are now many reports of α-amino acid peptide or protein-derived hydrogels that arise from the ordered aggregation of self-assembled fibers. (Aggelli 2003, Schneider 2002, Measey 2006, Mao 2004, Xu 2007) The highly ordered assembly of cyclic peptides has also been used to produce interesting structures analogous to pores. (Sanchez-Quesada 2002, Sanchez-Quesada 2001) On the abiotic side, an amphiphilic β-peptide foldamer was shown to self-assemble into a lyotropic phase. (Pomerantz 2006) There is also a growing list of reports involving abiotic foldamers that undergo a conformational change in response to some external stimulus (Figure 2.2). (Halder 2006, Balakrishnan 2006, Liu 2007, Khan 2006, Seebach 2006, Shin 2007)

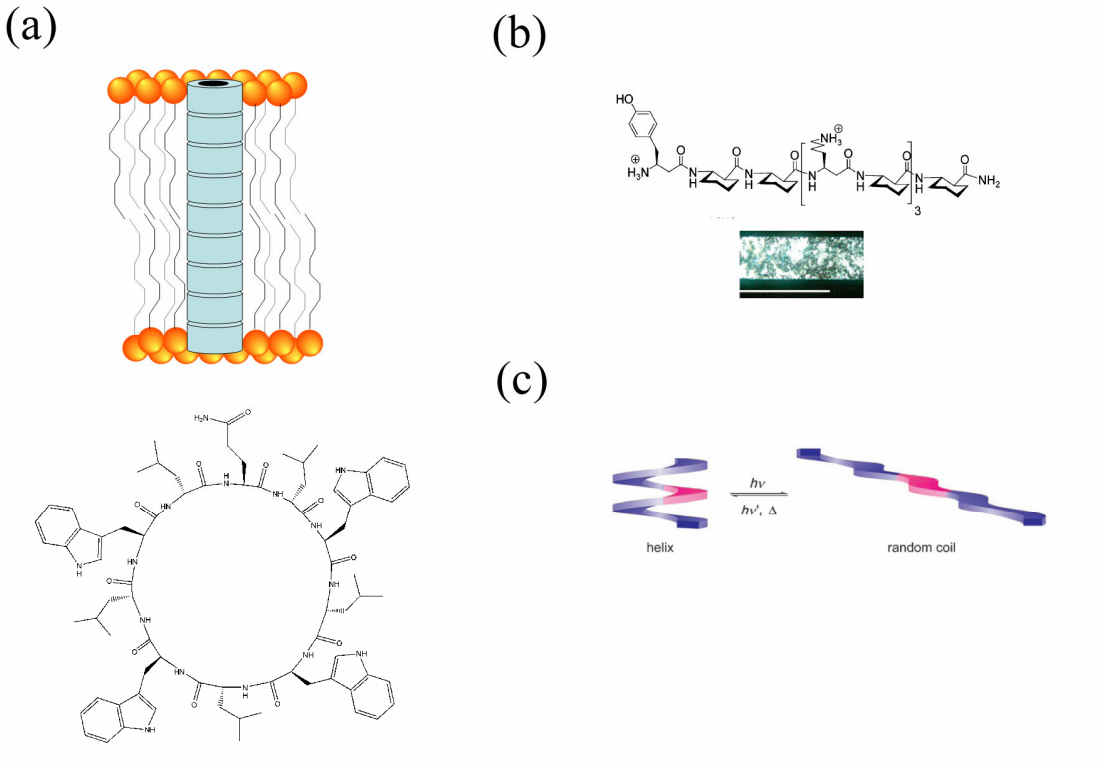


Figure 2.2 Highly ordered foldamers. (a) Synthetic cyclic peptide pores. (Sanchez-Quesada 2001) (b) Amphiphilic β-peptide foldamers in a lyotropic phase. (Pomerantz 2006) (c) Photoswitchable oligo(*meta*-phenylene ethylene) foldamer. (Khan 2006)

Many biological-derived hydrogels arise from aggregation of self-assembled fibers. Proteins, such as collagen, also form hydrogels via this pathway. (Mao 2004) Self-assembled and ordered protein fibrillar aggregates are symptomatic of an expanding list of diseases that now includes systemic amyloidoses (Merlini 2004), Alzheimer's disease (Finder 2007), Huntington's disease (Ross 2002), spongiform encephalopathies (Chakraborty 2005), Parkinson's disease (Cookson 2005), and Type II diabetes (Hoeppener 2006).

The common theme among the amyloid diseases is the presence of partially unfolded or misfolded proteins that self-assemble into highly ordered β-structured protofibrils followed by further assembly into amyloid fibrils (Figure 2.3). (Binder 2006) This conformational change and ordered assembly is irreversible, and the amyloid fibrils can form protein precipitates *in vivo*. Interestingly, in the case of Alzheimer's disease and transthyretin amyloidoses, there is growing evidence that the protofibrils are the toxic species rather than the mature amyloid precipitates. (Kayed 2003, Taylor 2003, Reixach 2004)

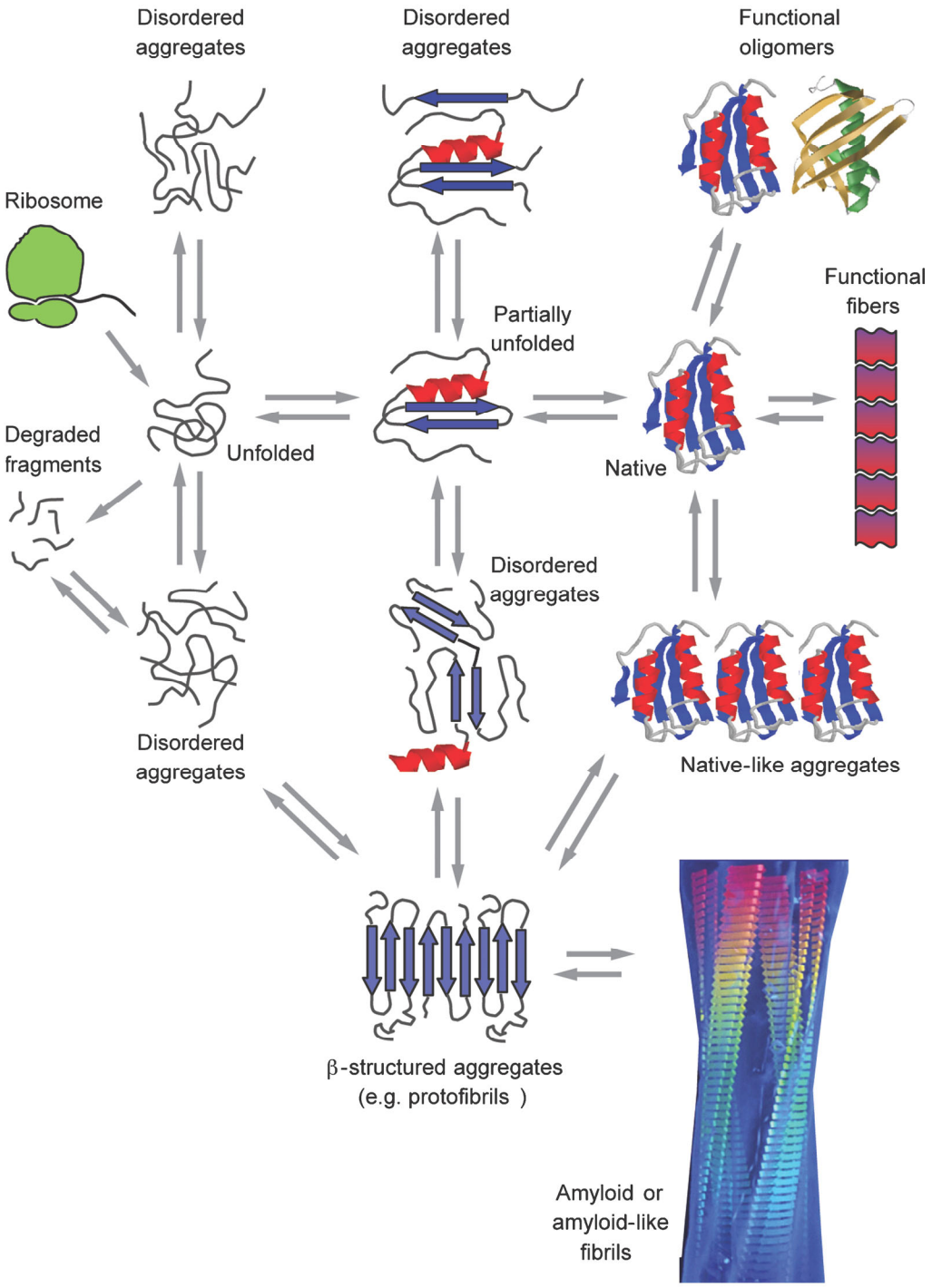


Figure 2.3 Schematic of the conformational states and pathways of proteins to the amyloid fibril. (Chiti 2006)

Many proteins have been shown to undergo the transformation to self-assembled fibrils. (Dobson 2003) The requisite cross-β fiber assembly is potentially accessible to all polypeptides by virtue of intrachain and interchain interactions of amide backbones. Although the main-chain interactions determine the overall structural theme of the amyloid, the side chain interactions of the specific polypeptide sequence determine the variations in the fibrillar structure. These structural variations due to polypeptide sequence coupled with structural variations from subtle changes in solution conditions, have led to the conclusion that while the cross-β fold and ordered self-assembly may be thermodynamically favored, the exact structural details are determined by kinetic accessibility. (Dobson 2003) There have been many systems designed to display this amyloid-like behavior, although almost all have utilized protein-derived and *de novo* designed peptides. (Hamley 2007, Teplow 1998, Makin 2005, Ciani 2002, Mimna 2007, Chen 2005, Matsumura 2002, Aggeli 2001, Wang 2005, Ray 2006, Deechongkit 2005, Mesquida 2007, Hirata 2007, Kammerer 2004)

Interestingly, recent studies have shown that the aggregation rates of polypeptides (unlike folding of the native structure) can be predicted using a simple "polymer" model. (Chiti 2003, Tartaglia 2004, Meinhardt 2007) The obvious difference between non-natural polymers and proteins is that proteins exhibit exquisitely well-defined native folds, which can persist in solution (without aggregation) because they represent local free energy minima as shown in Figure 2.4. The cross-β aggregates of these same polypeptides correspond to different, more global free energy minima with entirely different sets of ordered conformations. There are likely to be partial unfolding conformations lying adjacent to, or between, these ordered states on the free energy landscape. (Gregersen 2006)

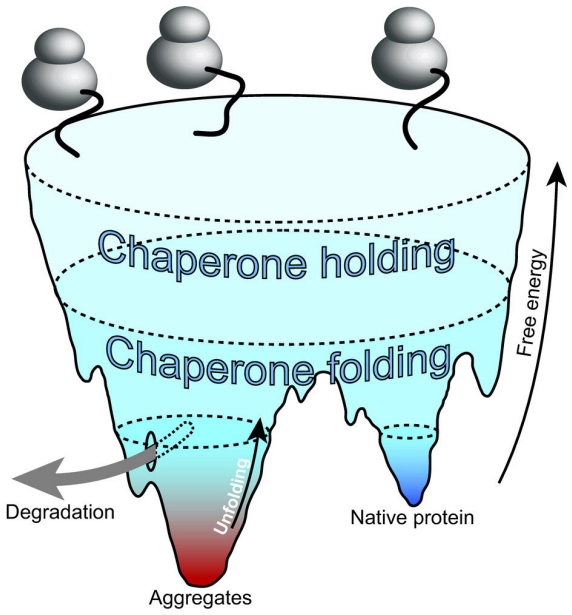


Figure 2.4 Free-energy folding landscape for chaperone-mediated protein folding. (Gergersen 2006)

The contrasting behavior of non-natural polymers and proteins leads to some interesting questions. First, is it possible to create an abiotic (i.e. non-polypeptide) system capable of exhibiting amyloid-like behavior involving two free energy minima; both of which exhibit a great deal of conformational order, the more thermodynamically stable of which is the ordered aggregate? If so, what structural characteristics of such an abiotic system modulate behavior and properties?

Previously, we reported an abiotic amphiphilic foldamer **2.1** that upon heating undergoes an irreversible conformational change leading to formation of a hydrogel. (Nguyen 1999) The amphiphilic foldamer **2.1** was constructed with linkers between aromatic units that contained one amino acid, which alternated between leucine and aspartic acid. The general structure created by alternating DAN and NDI units with amino acid linkers is termed *a*romatic *e*lectron *d*onor-*a*cceptor folda*mer*, or *aedamer*. When folded, the hydrophobic leucine residues are in position to reside on one side, and

the negatively charged aspartate residues are in position to reside on the opposite side of the pleated, stacked core as shown in Figure 2.5.

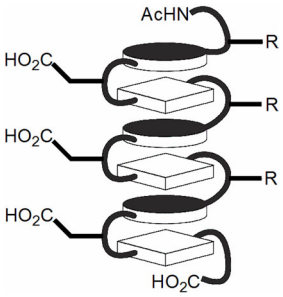


Figure 2.5 Cartoon of pleated, stacked structure of amphiphilic foldamer **2.1**.

Upon heating, amphiphilic aedamer **2.1** undergoes an irreversible transition to a hydrogel state with significant loss of the charge-transfer absorbance. A mechanism of hydrogelation was proposed that involved partial unfolding of the initial pleated structure followed by hydrophobically driven aggregation to a "tangled" hydrogel state in analogy to non-natural polymers as shown in Figure 2.6. The proposed mechanism took into account a well-defined, folded initial state, as any population of unfolded molecules should lead to hydrogel formation at room temperature (not observed). The "tangled" aggregate hydrogel final state, as originally described, was of unknown conformational order and its properties were not investigated. (Nguyen 1999) To the best of our knowledge, the previously reported amphiphilic foldamer **2.1** is the only foldamer reported to exhibit conformational switching combined with hydrogel formation behavior.

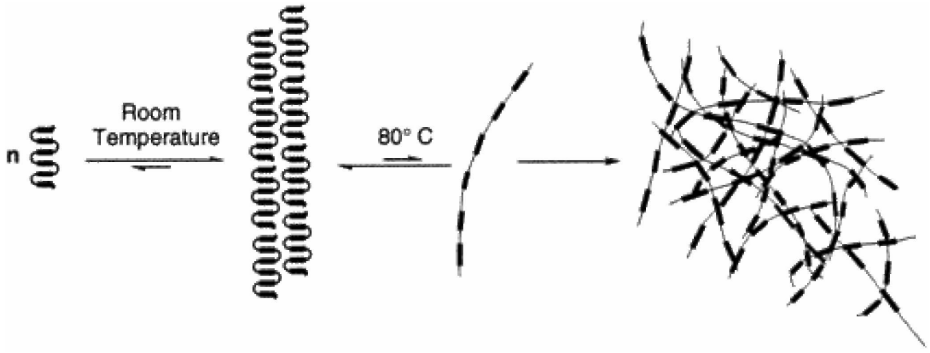


Figure 2.6 Original hydrogelation mechanism of **2.1** proposing unfolding of the pleated, stacked structure upon heating followed by "tangled" aggregation. (Nguyen 1999)

2.3 RESULTS

2.3.1 Aedamer Design and Synthesis

Figure 2.7 shows the series of aedamers synthesized to probe the effect of the hydrophobic linker amino acid residue on hydrogel properties of heated samples. In addition to the original leucine derivative, a valine, isoleucine, and norleucine derivative were synthesized in order to probe the influence of subtle structural changes on hydrogel properties. Note that compounds **2.1**-**2.3** are actually constitutional isomers.

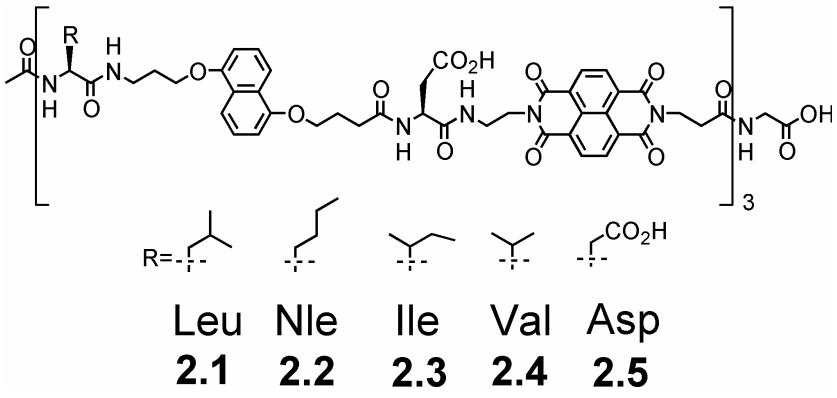


Figure 2.7 Structures of amphiphilic aedamers **2.1**-**2.5**.

All aedamer derivatives were synthesized using standard Fmoc solid phase peptide synthesis methods using monomer units previously reported. (Nguyen 1999) Experimental measures of the relative hydrophobicities of the aedamers **2.1**-**2.4** were investigated by a co-injection of the four compounds on reverse-phase HPLC as shown in Figure 2.8. The molecules eluted in the order **2.4** (Val), **2.1** (Leu), **2.2** (Nle), then **2.3** (Ile) indicating that the isoleucine derivative **2.3** was the most hydrophobic to a small degree, followed closely by the norleucine derivative **2.2** and leucine derivative **2.1**. Interestingly, valine compound **2.4** was significantly less hydrophobic than the other members of the series. To the best of our knowledge, this order of elution did not correspond exactly to any scale in the literature that includes norleucine. (Kovacs 2006, Tossi 2002)

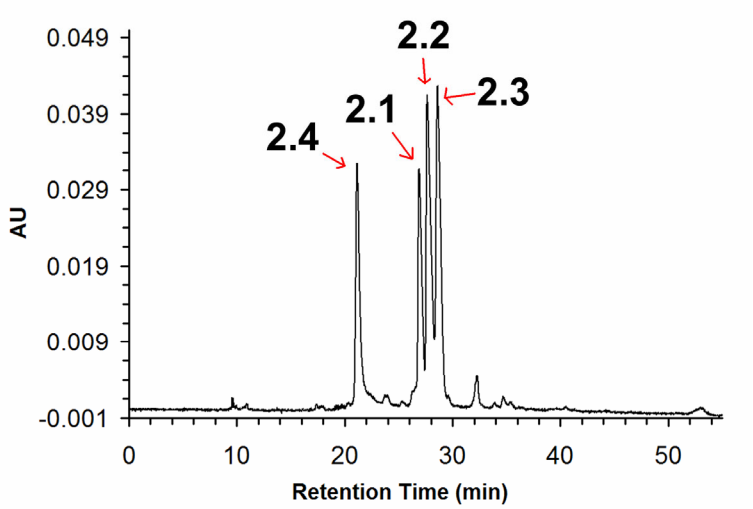


Figure 2.8 RP-HPLC chromatograph of co-injection of compounds **2.1**-**2.4**.

2.3.2 Light Scattering

Dynamic light scattering was measured for compounds **2.1**-**2.4** at 2.5 mM concentration in buffer prior to heating and analyzed assuming a globular protein structure and using the monomodal analysis of data provided by the manufacturer (Wyatt

Technology Corp.). The resulting estimated molecular weights for compounds **2.1**-**2.4** are 239 kDa, $1.65x10^3$ kDa, 210 kDa, and 340 kDa, respectively. Aggregation is expected for these gelling compounds as compound **2.1** was previously shown to be highly aggregated by NMR. (Nguyen 1999)

2.3.3 UV Spectroscopy

Face-centered stacking of the aromatic units of aedamers leads to pi molecular orbital overlap that causes a characteristic ~50% hypochromism in the 382 nm NDI absorbance. In order to quantify the hypochromism associated with folding, ultraviolet spectra (UV) were recorded for the series of aedamers (~15 μM) in aqueous buffer as well as in the presence of 2% (w/v) cetyl trimethylammonium bromide (CTAB) as shown in Figure 2.9. The positively charged CTAB detergent is known to unfold negatively charged aedamers by causing the unstacking of the aromatic units, so absorbance in CTAB is used to establish an absorbance value of unfolded aedamers for the hypochromism calculation. Observed hypochromism is reported as [(absorbance recorded in CTAB - absorbance recorded in buffer)/absorbance recorded in CTAB] x 100. The hypochromism values measured for compounds **2.1**-**2.5** are 51%, 52%, 52%, 53%, and 54%, respectively.

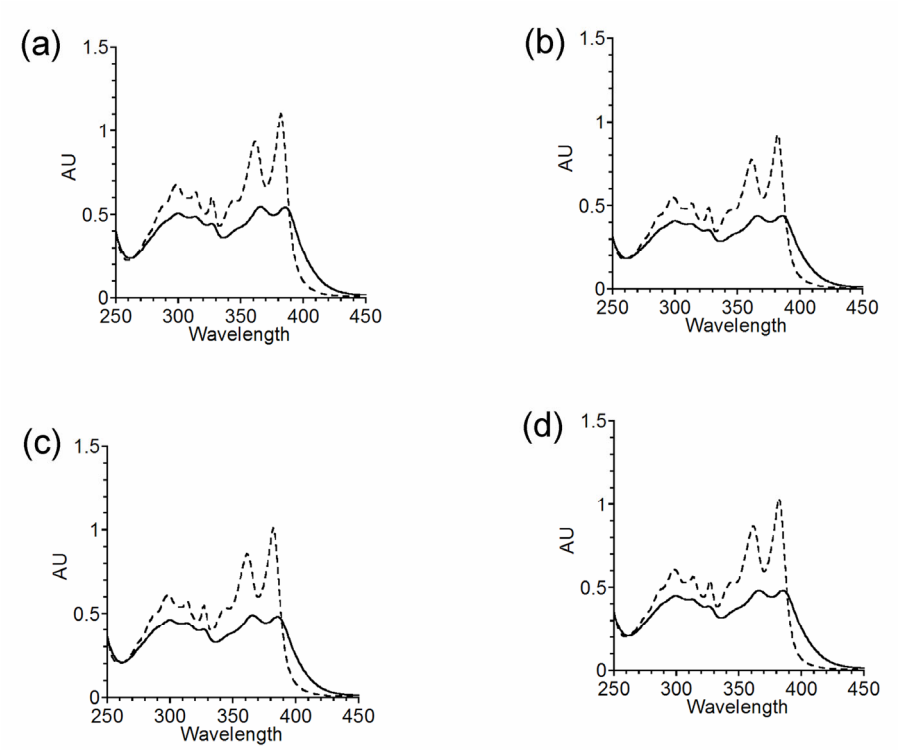


Figure 2.9 UV spectra of compounds **2.1**-**2.4** (a-d) with (---) and without (—) addition of 2% (w/v) CTAB indicating hyprochromism.

2.3.4 Hydrogelation

Upon heating 2.0 mM solutions (~ 0.50 wt%) of the compounds **2.1-2.4** at 80˚ C for 1.5 hours, compounds **2.1-2.3** formed viscous hydrogels. No hydrogel was seen in the case of the valine-containing 2.**4**. Compounds **2.1-2.3** were also found to be capable of forming hydrogels at concentrations at 1.0 mM (~0.25 wt%). Compound **2.4** did not produce a visible hydrogel even at its maximum solubility (~5 mM) after heating at 95˚ C for 2 hours.

2.3.5 Visible Spectroscopy

Figure 2.10 shows the change in the visible region charge-transfer absorbance for compounds **2.1**-**2.4** both before and after heating. The degree of charge-transfer

absorbance varied greatly with changes among the series. Compound **2.1** shows the smallest retention of charge-transfer absorbance after gelling, with only 28% of the original absorbance remaining. This loss in charge-transfer absorbance was accompanied by a shift in λ_{max} from 531 to 521 nm. Compound **2.2** retained 36% of the initial charge-transfer absorbance with a shift in λ_{max} from 529 to 511 nm. After accounting for light scattering, compound **2.3** retained 88% of the charge-transfer absorbance and a shift in λ_{max} from 525 to 502 nm. Unlike compounds **2.1**-**2.3**, the valine derivative **2.4** showed a 33% increase in the charge-transfer absorbance with a corresponding shift in λ_{max} from 518 nm to 514 nm. In all cases, the spectroscopic transitions that occurred upon heating were irreversible, even for compound **2.4**, as shown by repeated heating-cooling cycles.

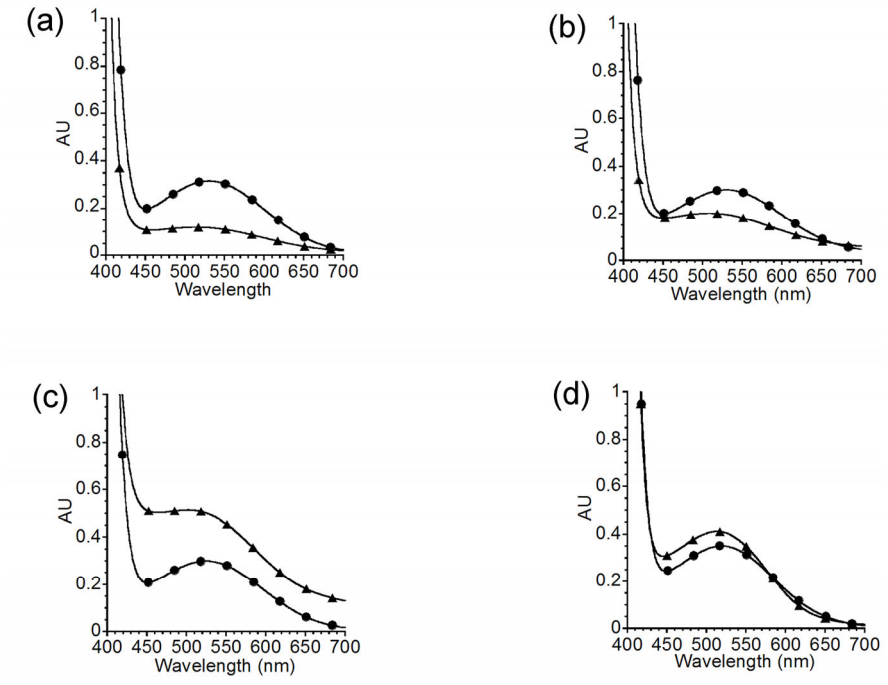


Figure 2.10 Visible spectra of compounds **2.1**-**2.4** (a-d) before (•) and after (▲) heating at 80° C.

2.3.6 Rheology

A variety of rheological experiments were performed to characterize compounds **2.1**-**2.4**, both before and after hydrogel formation. The absolute viscosity of a 4 mM solution of **2.1** was measured during heating from 25˚ C to 80˚ C, as shown in Figure 2.11. Consistent with aggregation indicated by the light scattering results, the viscosity at room temperature for the solution of **2.1** was considerably greater than that of pure water (0.365 Pa·s versus 0.001 Pa·s). Interestingly, the viscosity of the solution decreases initially during heating, even approaching that of water. Once 80˚ C is reached, the viscosity begins to increase rapidly corresponding to formation of hydrogel provided some pre-gelled material was placed in the rheology sample holder, which is presumably required because the rheology sample holder was entirely sealed and therefore did not allow for an air-water interface. Figure 2.11b displays an experiment where upon reaching 80° C, the viscosity is monitored as a function of time. While seeding was not used in this experiment, a rapid increase in viscosity is still seen. The decrease in viscosity after approximately five minutes of measurement is due to the constant motion of the rheometer, which disrupts the aggregates that are required to form the gel as they are being formed. This result indicates that the aggregation process is negatively impacted by mechanical forces. Interestingly, compound **2.4**, while not forming a hydrogel, did show some increase in viscosity, from 0.038 Pa·s prior to heating, to 0.143 Pa·s following heating at 80° C.

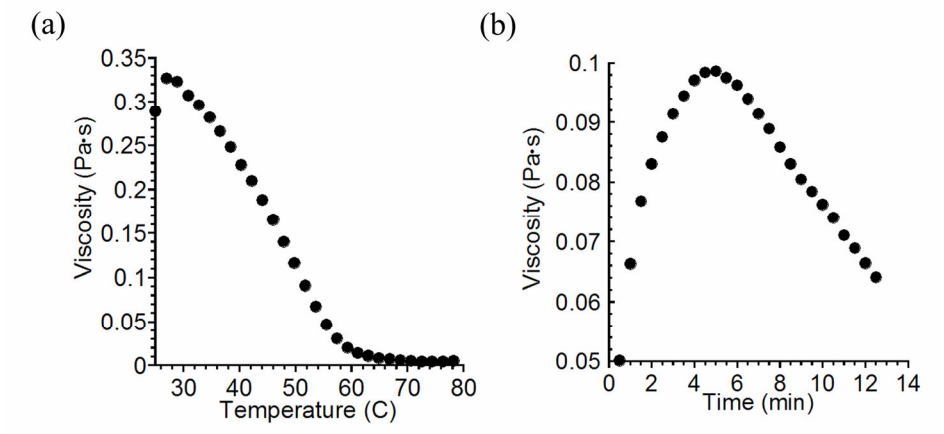


Figure 2.11 (a) Viscosity measured as a function of temperature during initial ramp to 80° C for compound **2.1**. (b) Viscosity measured as a function of time after reaching 80° C.

Hydrogels were analyzed following seeding samples of the different aedamers with 10% pre-gelled material. Without seeding, the gels formed in the rheometer were not reproducible enough for quantitative comparison. We attributed this lack of reproducibility to the importance of the air-water interface in the nucleation step of the self-assembly process, which was altered in the absence of seeding because of the required silicon oil layer used to prevent evaporation in the rheometer.

Hydrogel behavior is indicated by a frequency-independent storage modulus (G'; a measure of elastic behavior) and loss modulus (G"; a measure of liquid behavior, i.e. the ability of the sample to flow). Figure 2.12a depicts a frequency sweep experiment in the linear viscoelastic regime at constant strain of 0.5% for compounds **2.1**-**2.3** at 4 mM concentration (~1wt%) after heating. Such frequency independence is consistent with gel behavior. (Gupta 2000) Both the storage and loss moduli of compound **2.4** were not frequency independent, as they both increased at higher frequency as shown in Figure 2.12b. For a direct comparison of elastic gel strength, the equilibrium storage moduli for

compounds **2.1**-**2.4** at a frequency of 2.64 rad/s are 323, 652, 1400, and 0.22 Pa, respectively.

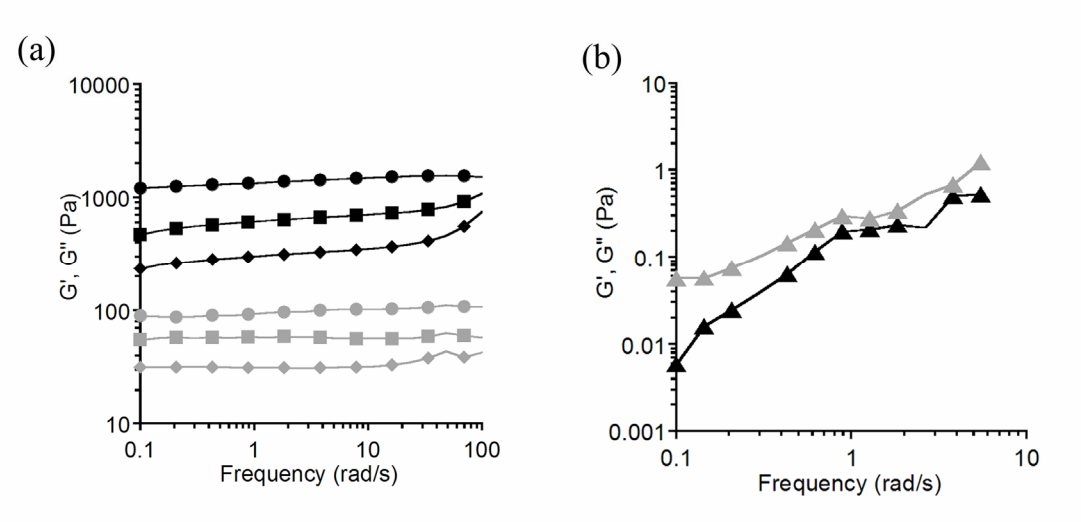


Figure 2.12 Frequency sweep data of storage (G') and loss (G") moduli for (a) compounds **2.1** (G'= ♦, G"= ♦), **2.2** (G'= ■, G"= ■), **2.3** (G'= ●, G"= ●) and (b) compound **2.4** (G'= ▲, G"= ▲).

Another characteristic feature of hydrogel behavior is a larger storage modulus versus loss modulus (G'>G"). Values of the loss tangent, tanδ = G"/G', at a frequency of 2.64 rad/s for compounds **2.1**, **2.2**, and **2.3** are 0.096, 0.089, and 0.069, respectively. However, in the case of the valine derivative compound **2.4**, the loss modulus was found to be larger (G">G') over the frequency range investigated. Again, these data are consistent with compounds **2.1**, **2.2**, and **2.3** exhibiting hydrogel behavior, while compound **2.4** does not.

Self-assembled hydrogels exhibit shear-thinning, that is a decline in absolute viscosity with an increase in shear rate due to disruptions of non-covalent cross-links within the gel. (Schneider 2002, Terech 1998) Figure 2.13 shows the representative shear-thinning behavior of hydrogels of compounds **2.1**, **2.2**, and **2.3**.

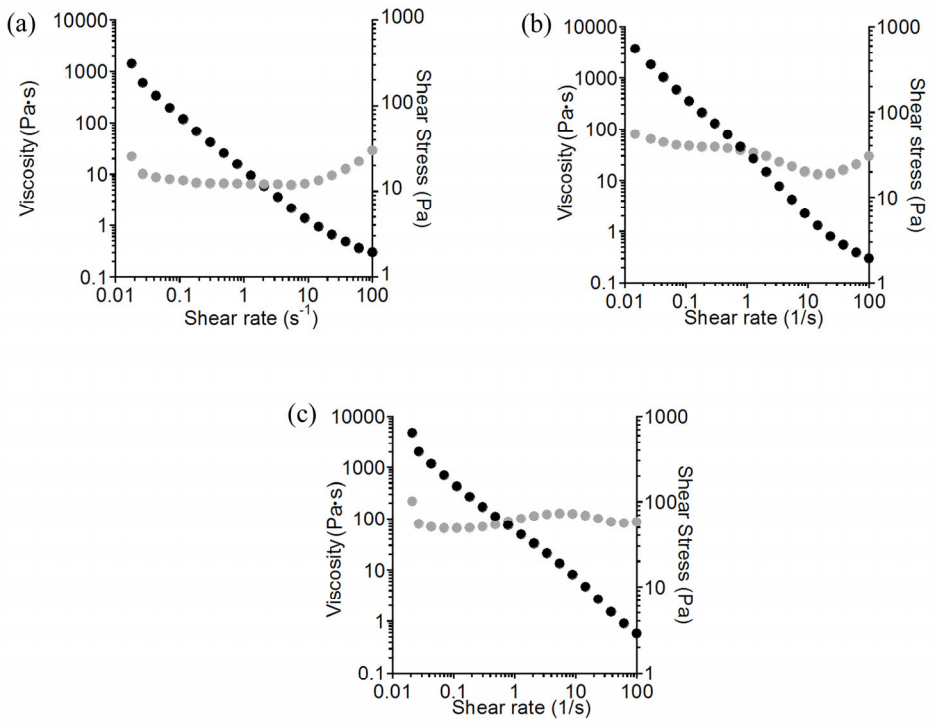


Figure 2.13 Shear thinning behavior of compounds **2.1**-**2.3** (a-c) after heating to 80° C.

Figure 2.14a shows a time sweep experiment at 25° C where, after the gel formation of compound **2.1** is complete, 1000% strain is applied for 180 s, the strain was removed, and the storage and loss moduli were monitored. The hydrogel of compound **2.1** regained 94% of its initial elastic strength after only 15 minutes. The time sweep experiments of compounds **2.2** and **2.3** are also given in Figure 2.14.

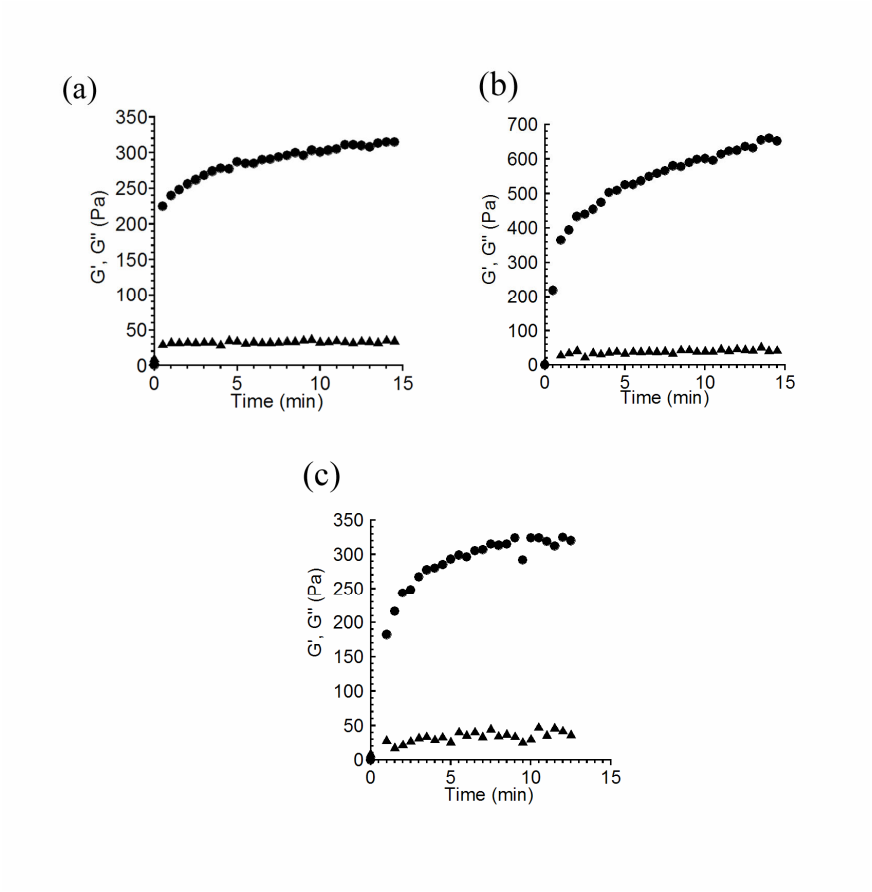


Figure 2.14 Recovery of gel strength of compounds **2.1**-**2.3** (a-c) after 1000% strain was applied for 180 s.

2.3.7 Circular Dichroism

As a probe of possible higher order structure in the hydrogel state, CD spectra of the non-gelled and hydrogel form of compounds **2.1**-**2.4** were compared as shown in Figure 2.15. In addition to compounds **2.1**-**2.4**, a non-amphiphilic aedamer with all aspartate residues (compound **2.5**) was used as an important control molecule with well characterized folding in aqueous solution, yet no ability to form a hydrogel state. As expected, the control compound **2.5** showed a relatively small signal in the carbonyl

region (<250 nm) and nothing in the aromatic region (310 nm – 420 nm). There was no change in its CD spectrum observed upon heating.

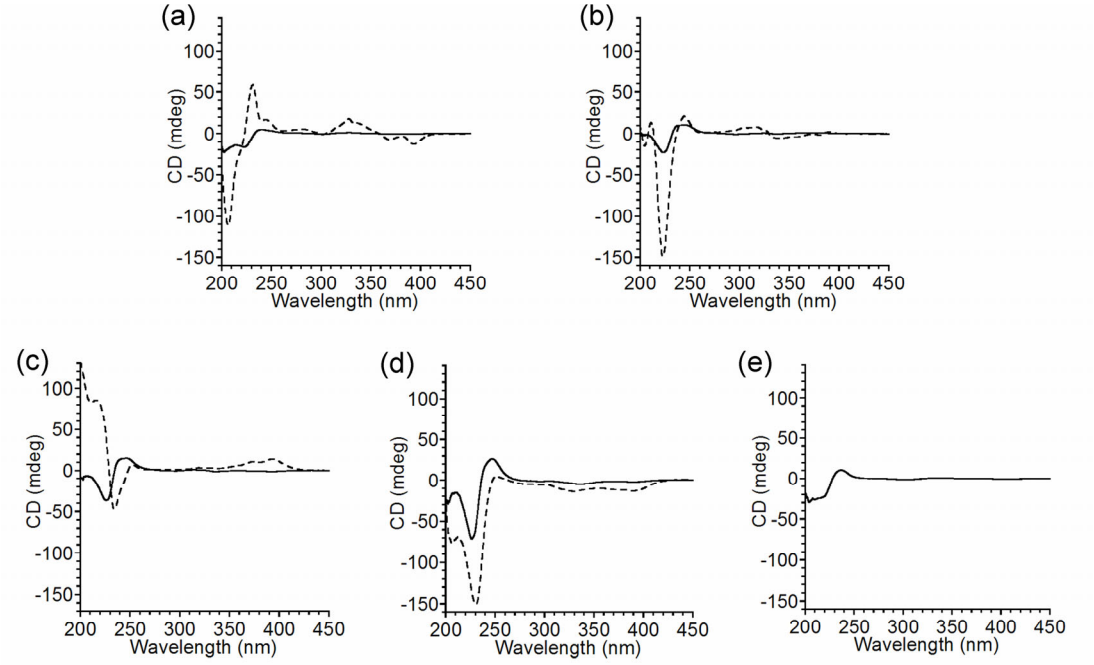


Figure 2.15 Circular dichroism spectra of aedamers at 0.2 mM concentration in 50 mM sodium phosphate pH 7, 100 mM NaCl before heating (—) and after heating (---) at 80° C for 1 hour. (a) **2.1** (b) **2.2** (c) **2.3** (d) **2.4** (e) **2.5**

Before heating, compound **2.1** gave a CD spectrum with only a very weak signal in the carbonyl region, similar in shape and intensity to the control compound **2.5**. After heating to the hydrogel state, a strong negative Cotton effect, centered at 206 nm, was observed followed by a strong positive Cotton effect centered at 231 nm. There are also less intense positive followed by negative Cotton effects in the aromatic region, centered at 328 nm and 393 nm, respectively. Compound **2.2** exhibited a strong negative Cotton effect at 223 nm with a relatively small maximum and minimum in the aromatic region at 316 nm and 342 nm. Compound **2.3** showed a strong positive Cotton effect at the lowest

wavelength measured (200 nm) with a maximum at 217 nm and a minimum at 234 nm in the carbonyl region as well as a positive Cotton effect in the aromatic region centered at 393 nm.

The CD analysis of compound **2.4** turned out to be somewhat surprising. The unheated sample of **2.4** was seen to have the strongest CD signal of any unheated sample, with negative Cotton effects observed at 205 nm and 231nm in the carbonyl region, but also visible are small negative Cotton effects at 329 nm and 391 nm in the aromatic region. Despite the fact that this compound did not form a hydrogel, the negative Cotton effect CD signals present in the unheated sample grew significantly following heating at 80°C.

Interestingly, compound **2.4** showed the same more intense negative Cotton effects when incubated at 25˚ C for 2 weeks as was seen following heating at 80° C. On the other hand, samples of compounds **2.1**-**2.3** showed little change when incubated at 25° C for 2 weeks.

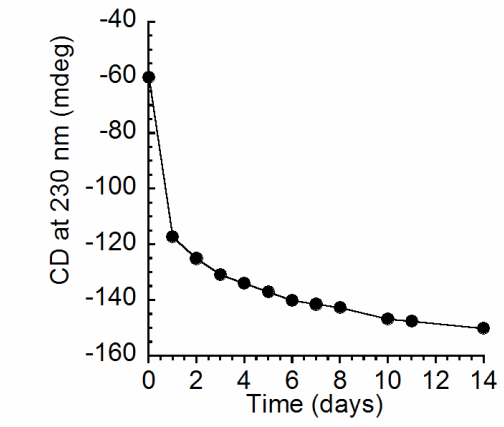


Figure 2.16 Time dependence of CD signal of compound **2.4** at 230 nm.

While nucleation with pre-gelled material enhanced the reproducibility of gel formation in the absence of an air-water interface in the rheology experiments, circular

dichroism experiments were designed to determine if the addition of pre-gelled material could catalyze the conformational switching process at lower temperatures. The CD spectra of compounds **2.1**-**2.3** before and after heating to 80° C were compared to the spectra of samples incubated with 10% pre-gelled material at 25° C and 37° C for two weeks as shown in Figure 2.17. The leucine compound **2.1** showed intermediate CD spectra at 25° C and 27° C compared to the unheated and 80° C spectra. Compounds **2.2** and **2.3**, however, showed very different spectra at 25° C and 37° C compared to the unheated and 80° C samples.

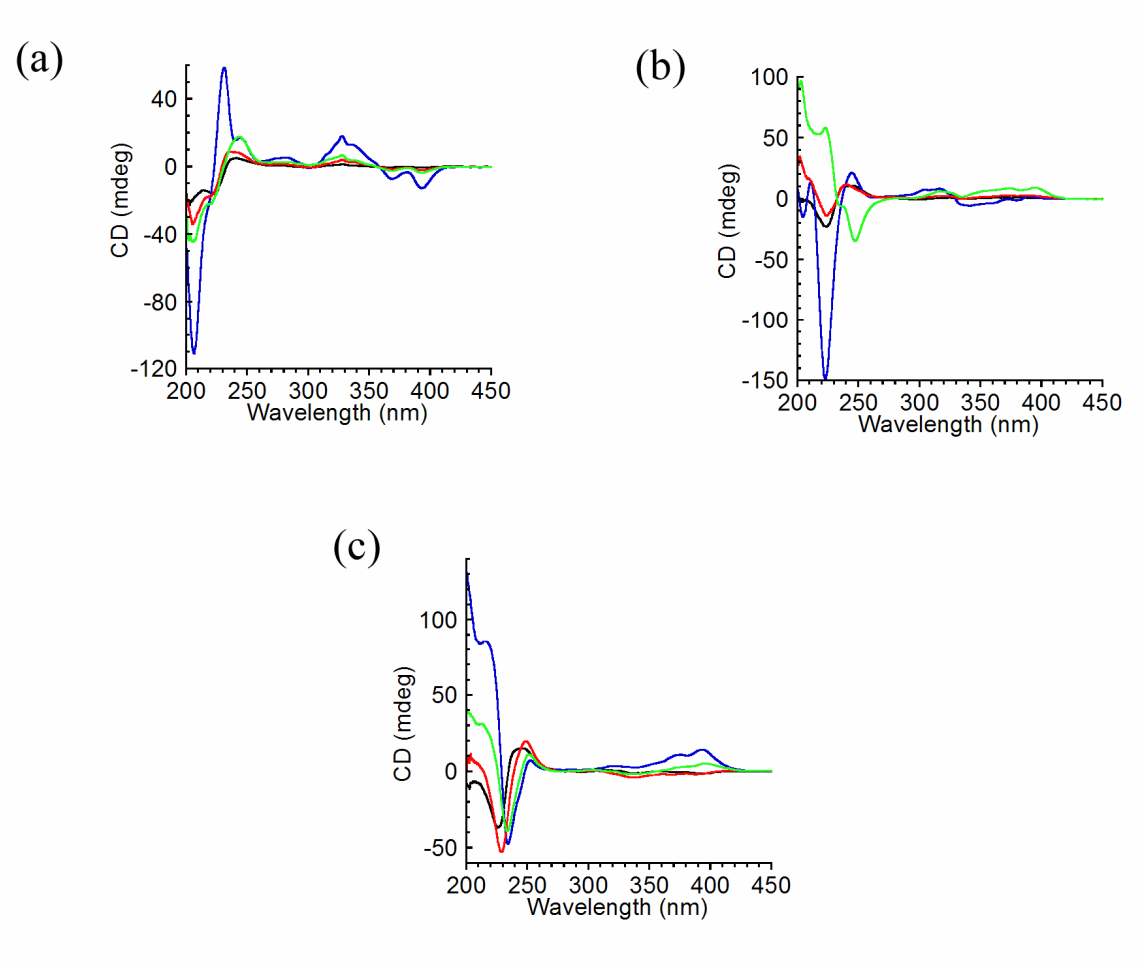


Figure 2.17 CD spectra of compounds **2.1**-**2.3** (a-c) before heating (—), after heating to 80° C (—), after 2 weeks at 25° C with 10% pregelled material (—), after 2 weeks at 37° C with 10% pregelled material (—).

2.3.8 Microscopy

Compound **2.1**-**2.3** hydrogel morphology was probed via SEM of desalted, lyophilized samples as shown in Figure 2.18. The SEM image of compound **2.1** after heating in Figure 2.18a showed highly branched fibers that form a three-dimensional mesh-like scaffold. Compounds **2.2** and **2.3** showed similar morphologies after gelling, with compound **2.3** displaying thicker fibers. The SEM of non-gelling compound **2.4** appeared as a large mass of material with no discernable fine structure similar to the images of compounds **2.1**-**2.3** prior to heating.

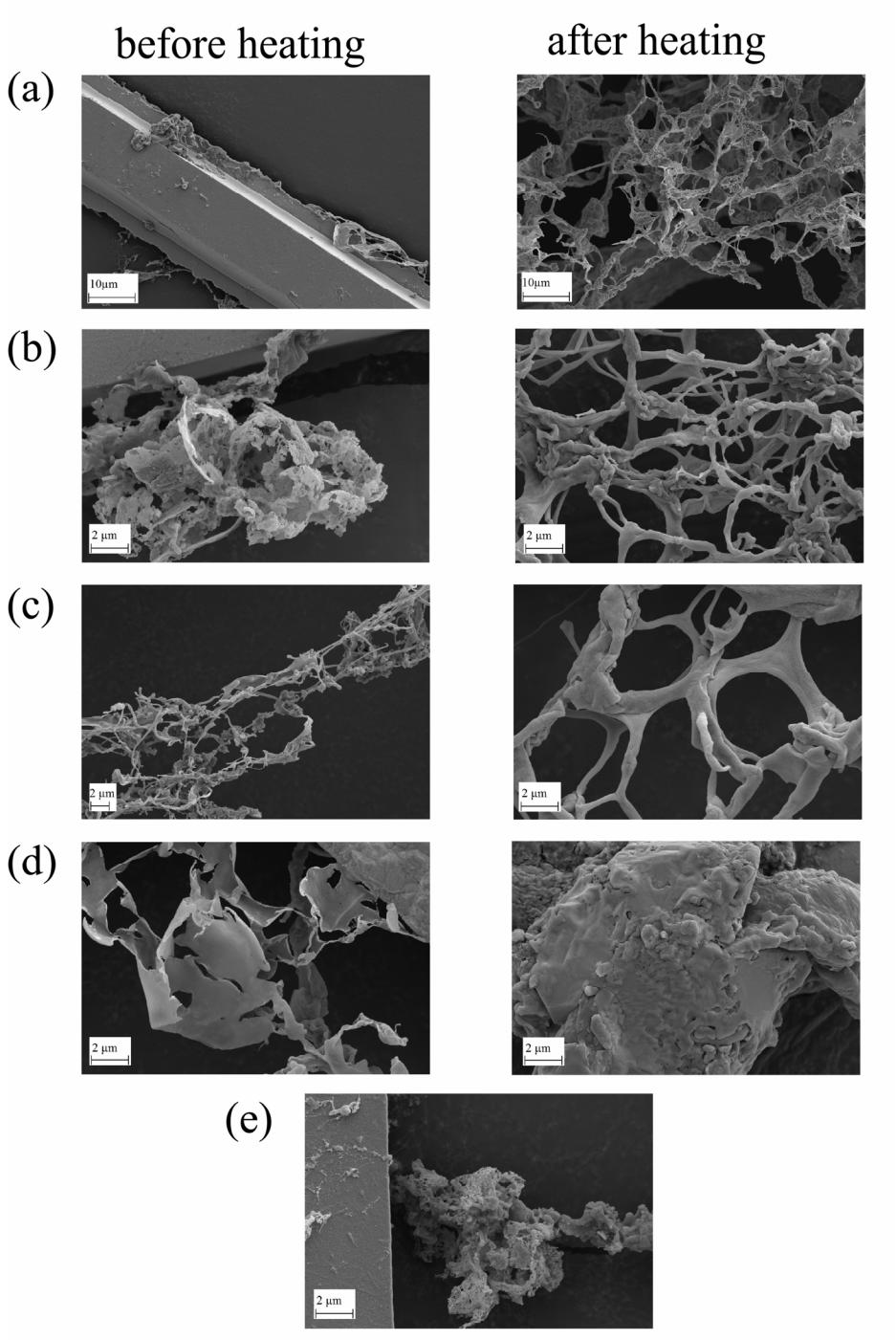


Figure 2.18 SEM images of gel morphology from lyophilized copper grids before (left) and after (right) heating to 80° C for compounds (a) **2.1**, (b) **2.2**, (c) **2.3** (d) **2.4**, (e) **2.5**.

2.4 DISCUSSION

2.4.1 Similar Folding Prior to Heating

Quantitative analysis of the spectroscopic properties of unheated samples of the amphiphilic aedamers **2.1**-**2.4** is consistent with complete folding analogous to our previously reported aedamer designs. In particular, hypochromism arises from the face-centered stacked geometry of the DAN and NDI aromatic units. The extent of hypochromism is dependent on distance r as a function of $1/r^3$ and is highly orientation dependent with respect to the angle formed by the aromatic ring planes. (Lokey 1995) Prior to heating, compounds **2.1**-**2.4** exhibited nearly identical levels of hypochromism, between 51-53%. This similarity of hypochromism to that seen for the well-studied aedamer **2.5** provides strong evidence that all the amphiphilic aedamers used in this study are not only folded, they are folded to the same extent and with the same pleated, stacked conformation. For comparison, formation of the DNA double helix from DNA single strands results in an observed hypochromism of roughly 40%, the exact value of which is dependent on base sequence. (Weissbluth 1971) As further support for the folding of compounds **2.1**-**2.4**, it should be noted that they all exhibit the intense purple charge transfer absorbance (similar to the color of a fine merlot) that is characteristic of DAN-NDI stacking in water.

2.4.2 Aggregation and Hydrophobicity

A major difference between the amphiphilic aedamers **2.1**-**2.4** and the control aedamer **2.5** is that the amphiphilic aedamers aggregate in solution, as evidenced by significant viscosities prior to heating as well as their observed dynamic light scattering. Based on the dynamic light scattering data, the estimated molecular weights of the aggregates at 2.5 mM concentration are on the order of 90-400 molecules per aggregate.

The control aedamer **2.5**, bearing 7 negative charges and no hydrophobic side chains, is monomeric in solution under these conditions. (Zych 2000) Unfortunately, the aggregation prevents any detailed NMR conformational analysis of the stacked geometries of compounds **2.1**-**2.4** due to signal broadening.

The linker residues for compounds **2.1**-**2.4** were specifically chosen for their subtle geometrical differences. As a qualitative experimental measure of overall hydrophobicity, the relative retention times were compared on reversed-phase HPLC. Compounds **2.1**-**2.4** elute in the order **2.4** first, followed by a gap, then in quick succession; **2.1**, **2.2**, **2.3**. Considering a relatively slow gradient was used (gradient increase of only 0.36% acetonitrile per minute), these differences in retention time are relatively small, consistent with the notion that overall relative molecular hydrophobicity is similar for compounds **2.1**-**2.4** (at least as it pertains to interactions with the C-18 HPLC chromatography support).

2.4.3 Hydrogel Formation

Analysis of compounds **2.1**-**2.5** after heating at 80 ˚C shows that compounds **2.1**-**2.3** form self-supporting hydrogels with varying degrees of the characteristic purple color remaining. Surprisingly, despite having an amphiphilic design similar to **2.1**-**2.3**, compound **2.4** did not show gelling behavior under these conditions. Control compound **2.5**, with only aspartic acid side chains, also showed no gelling behavior, consistent with expectations that assume amphiphilic character is a prerequisite for hydrogel formation.

The significant reduction in the charge-transfer absorbance upon hydrogel formation of compound **2.1** supported the original model of a “tangled” (non-ordered) state within the hydrogel, because it was assumed that any order in the hydrogel would necessarily derive from systematic donor-acceptor aromatic stacking. Relevant to the present study, an entirely non-ordered, random mode of aggregation in the hydrogel

would have a statistical distribution of aromatic donor-acceptor interactions producing an average charge-transfer band that should remain relatively consistent among our series of amphiphilic aedamers. However, within the series of compounds **2.1**-**2.3**, surprising variation was seen in the retention of the charge-transfer absorbance upon hydrogel formation, ranging from 28% retention of the original charge transfer absorbance intensity for **2.1** to 88% retention for **2.3**. This observed substantial variation indicates that differences in side chain structure are serving to strongly modulate either inter- or intra-strand aromatic donor-acceptor stacking interactions in the hydrogel state. Given the relatively subtle differences in side chain structures for **2.1**-**2.3**, it would be surprising if such large differences in aromatic stacking would be manifest in entirely unordered, tangled aggregates as originally conceived.

Viscoelastic experiments provided a more complete understanding of the material properties of compounds **2.1**-**2.3** as well as insight into the hydrogel assembly process. Compounds **2.1**-**2.3** have the characteristic features of a hydrogel: frequency-independent moduli, storage modulus greater than loss modulus, and shear thinning behavior. Although compounds **2.1**-**2.3** share these characteristic features, there are significant differences in elastic behavior within the series. As a larger equilibrium storage modulus indicates a stronger, more elastic hydrogel, the trend within the series of compounds **2.1**-**2.3** indicates the isoleucine-containing aedamer **2.3** produces the strongest, most elastic hydrogel, followed by the norleucine-containing aedamer **2.2**, and leucine-containing aedamer **2.3**, the weakest.

A property of self-assembled gels that a polymer tangled aggregate does not share is the ability to quickly recover elastic strength after a period of strain is applied to the material. (Chen 2004, Nowak 2002) While compound **2.1** gave reproducible results in this experiment, the graphs of the recovery of compounds **2.2** and **2.3** had similar shapes

but the endpoint of the recovery was not reproducible. In some cases, the endpoint of recovery gave a larger storage modulus than before the strain was applied, and in others, the endpoint was much less than the starting storage modulus. The hydrogel of compound **2.1** regained 94% of its initial elastic strength after only 15 minutes. This result further suggests the original non-ordered, tangled model proposed for the hydrogel self-assembly of **2.1** is likely not correct. In addition to the quick recovery of the gel strength, the size of the aedamers must be taken into account. Polymers often have molecular weights that are at least an order of magnitude larger than compounds **2.1**-**2.4** (2.7 kD), which have a size similar to a 20-25 residue peptide. It is improbable that such small molecules forming a non-ordered tangled aggregate could hold water in a hydrogel at concentrations as low as 0.5 wt%.

Hydrogel assembly apparently begins with the disruption of the room-temperature intramolecular aggregation of folded molecules as shown by an initial decrease in viscosity upon heating. Following this loss of viscosity, hydrogel self-assembly begins (viscosity rises sharply), presumably mediated by molecules that are unfolded to some degree. The self-assembly process is apparently nucleation dependent as indicated by requirements for either 1) an air-water interface in the hydrogel chamber, or 2) the seeding with a small amount of hydrogel in the absence of an air-water interface. Importantly, for **2.1**-**2.3**, hydrogel formation is accompanied by unfolding of the original pleated, stacked folded structure as evidenced by loss of the charge transfer absorbance to varying degrees. Thus, it appears that for these derivatives, hydrogel formation involves conversion to an alternative conformation.

2.4.4 Highly Ordered Aggregation

The CD spectra recorded for **2.1**-**2.3** prior to heating show that the absorbing chromophores are not in a highly ordered chiral environment. In particular, the CD

spectra recorded prior to heating in all cases showed minimal signal intensity, with similar peak shapes among the series, including control compound **2.5**, and the non-hydrogel forming amphiphilic aedamer **2.4**. The folded aedamer solution structure does not have obvious structural features that would be expected to give rise to strong CD signals. In particular, as opposed to a protein α-helix or β-sheet structure, the carbonyl groups in the folded aedamer backbone are relatively mobile, far apart from each other, and they interact mainly with aqueous solvent. Therefore, folded aedamer backbone carbonyls would be expected to have only minimal coupled-oscillator interactions (that are chiral) with each other, consistent with the observed relatively small signals in the far-UV region of the CD. Previous studies have indicated that stacking of the alternating aromatic units in the aedamer core involves a perpendicular arrangement of the aromatic long axes. In other words, the resulting high symmetry of the stacked core is not strongly influenced by the chiral centers of the linking chains, so again, only a small CD signal is expected in the aromatic region, consistent with what was seen.

The CD spectra for compounds **2.1**-**2.3** in the hydrogel state show significantly enhanced signals, with features that are qualitatively very different for the different compounds. The most straightforward interpretation of these observations is that hydrogel formation involves a mode of intermolecular aggregation that is highly ordered, generating a strongly chiral environment around the various chromophores, especially those in the far UV region. The linker backbone carbonyls likely make a major contribution to the far UV signals, although it should be kept in mind that absorbances centered on the aromatic units, including the carbonyls of NDI, may be making a contribution to these far-UV signals as well. The substantial differences in CD signal shapes indicate that the different compounds produce different aggregated species, despite their overall similar molecular structures and original folded conformations.

Interestingly, the CD of samples seeded with 10% pregelled material and incubated at 25° C and 37° C gave different, non-intermediary spectra for compounds **2.2** and **2.3**. This result suggested that the conformational switching behavior is not necessarily a single transition between two states, but rather a transition through multiple partially or fully unfolded states to the final conformation in the aggregated state. The results of compound **2.1** showed intermediate spectra at 25° C and 37° C, which is more typical of a single transition between two states similar to valine compound **2.4**.

2.4.5 Comparison to Amyloid

The hydrogels formed from **2.1**-**2.3** can be compared to amyloid fibril formation by natural proteins. In each case, the molecule can adopt either a compact folded structure, or it can be converted to a highly ordered aggregate. In the case of amyloid-forming proteins, the highly ordered aggregate state is composed of amphiphilic β-sheet structures, based on an alternating hydrophobic-hydrophilic side chain pattern. In the case of the aedamers, the exact nature of the hydrogel forming aggregate is unknown at this time. An attempt to identify infrared (IR) signatures consistent with strong hydrogen bonds involving the backbone linker carbonyl groups in the hydrogel versus solution state failed to give unambiguous results for **2.1**-**2.3** due to overlap of the diimide signatures and experimental setups. Formation of a hydrogel in a ZnSe flow cell at 80° C was unsuccessful as heating caused bubble formation in the window. ATR-IR was attempted as well. Application of a 4 mM hydrogel directly to the diamond ATR crystal gave no peaks. Using a stream of air, the water was evaporated to give a hydrated thin film of the hydrogel. This method gave some peaks however, the signal-to-noise was not sufficient to give meaningful results. The use of a larger ATR crystal may provide better signal-to-noise. What we do know is that the different compounds produced hydrogels with significantly differing amounts of DAN-NDI stacking, as indicated by the differing

intensities of the visible charge transfer bands present in each. It is not clear whether these charge transfer absorbances derive from intramolecular or intermolecular stacking, or a mixture of both, so presently we are unsure of the extent to which DAN-NDI stacking is involved in the assembly process. The bottom line, however, is that like amyloid forming proteins, the CD measurements have indicated that amphiphilic aedamers **2.1**-**2.3** irreversibly form an insoluble, highly ordered aggregate. To the best of our knowledge, the amphiphilic aedamers **2.1**-**2.3** are the first folding abiotic molecules to undergo such a transition.

2.4.6 Branched Fibrils Lead to Hydrogel Formation

Microscopy was used to investigate the structural origin of the hydrogel properties of compounds **2.1**-**2.3**. The SEM micrographs of the hydrogels reveal networks of fibers surrounding relatively large cavities. The difference in gel strength and elasticity measured by rheology correlate with the thickness of the fibers in the SEM images as is most obvious when comparing the hydrogel SEM images of the norleucine (**2.2**) and isoleucine (**2.3**) derivatives. Stronger hydrogels are derived from thicker fibers. The structures of the hydrogel forming fibers seen in the SEM images for all three are branched (molecular networks) and larger in diameter than most proteinaceous amyloid fibrils, which do not generally form network. Apparently, the ability of heated samples of compound **2.1**-**2.3** to hold water as hydrogels arises from their molecular networks, whereas the proteinaceous amyloids more often form insoluble precipitates.

2.4.7 Valine Derivative

Prior to heating, the UV-Vis spectra of the valine compound **2.4** is virtually identical to that of **2.1**-**2.3** indicating similar folding in solution. However, following heating, **2.4** does not form a hydrogel like **2.1**-**2.3**. The CD spectra taken before and after

heating indicates that heating extends the order within samples of **2.4**, and this increase in order appears to be irreversible, like the others. As opposed to the other three, however, high temperature is not required to increase order, as incubation of **2.4** at room temperature for an extended period leads to the same increase in order seen following heating at 80°C. The SEM image of heated material reveals that **2.4** does not form a fibrillar network like the other three. This may be a reflection of lower association energies between molecules of **2.4** following heating, and/or perhaps a different geometry of association that does not lead to fibrillar networks, which are presumably required for hydrogel formation. The bottom line is that the subtle change in chemical structure of **2.4**, having one -CH_2- group less in each of the three hydrophobic side chains as compared with compounds **2.1**-**2.3**, is enough to have a dramatic influence on the properties of the aggregate formed after heating. Therefore, it appears that the appearance of hydrogel properties is exquisitely sensitive to the nature and probably geometry of intermolecular aggregation that occurs upon heating.

Interestingly, both the UV-Vis spectra and CD measurements indicate that **2.4** does not rearrange its core stacking as it produces a more ordered aggregate, since the charge transfer band remains strong and the CD spectral features maintain the same appearance, but are amplified, with heating or when left at room temperature for extended periods. Compounds **2.1**-**2.3** show various degrees of loss of the charge transfer band, require heating, and have altered CD spectral features following conversion to the more ordered aggregate following heating. A reasonable conclusion is that **2.1**-**2.3** undergo more significant core stacking reorganizations when the initially folded state transitions to the ordered, hydrogel-forming aggregate. This requirement for an alternative conformation may explain why **2.1**-**2.3** must be heated to produce the ordered aggregate, while **2.4** does not.

2.4.8 Kinetic and Thermodynamic Considerations

In all cases, **2.1**-**2.4**, the common behavioral feature is that initial aggregation of the folded aedamers in solution gives way to the irreversible formation of a more highly ordered aggregate upon heating (or just standing in the case of **2.4**). The irreversible nature of this transition could be a reflection of either kinetic or thermodynamics effects, or a combination of both. If the highly ordered state is the thermodynamically most stable form of the material, it is reasonable to propose that some sort of cooperative interactions between chains occurs that overcomes the significant entropic barrier imposed by ordered assembly.

2.4.9 Subtle Structural Differences Lead to Large Differences in Properties

Having only three examples of molecules capable of forming hydrogels makes any correlation between structure and properties preliminary. Nevertheless, interesting trends were observed for **2.1**-**2.3**. In the order **2.1** then **2.2** then **2.3**, retention of the charge-transfer absorbance in the hydrogel state increases, as does elastic hydrogel strength. It is tempting to propose that the increased residual charge transfer absorbance in the hydrogel formed from **2.3** is providing stronger associations between strands, but we cannot tell the difference between intramolecular and intermolecular stacking so such speculation is risky.

The hydrogel properties trend is also mirrored in the HPLC retention times on a C-18 column, in that **2.1** eluted first, then **2.2**, and **2.3** eluted last. To the extent that the HPLC retention times reflect relative molecular hydrophobicities, the implication is that increasing linker side chain hydrophobicity leads to enhanced interactions between molecules in the hydrogel state, thus offering another possible explanation for the increased elastic strength of the gels formed from **2.3**, with **2.2** and **2.1** having intermediate and the least elastic strength, respectively.

The relative HPLC retention times can be compared to the calculated hydrophobicities for the different side chains. Using a simple surface calculation (Spartan 04, side chains only) gives hydrophobic surface area values for valine (112 $Å^2$), isoleucine (132 $Å^2$), leucine (133 $Å^2$), and norleucine (137 $Å^2$). Using a semi-empirical method based on accessibility of hydrophobic surface area within protein structures, Karplus reports values for valine (135 $Å^2$), isoleucine (155 $Å^2$), and leucine (163 $Å^2$). (Karplus 1997) Using this same approach, we come up with a value for norleucine (165 $Å^2$) that is very similar to leucine. Using either approach, valine is less hydrophobic by a wider margin, as isoleucine, leucine, and norleucine are relatively similar. This trend can be used to explain the reversed phase HPLC elution order of **2.4** (valine) followed by a gap then **2.1** (leucine), **2.2** (norleucine), then **2.3** (isoleucine) in quick succession. Note that the elution order of the **2.1**, **2.2**, and **2.3** does not exactly follow from calculated hydrophobic surface area considerations alone.

It is important to keep in perspective the relative differences seen with **2.1**-**2.3** before and after heating. The structural differences between the molecules are subtle, since **2.1**-**2.3** are in reality constitutional isomers. Consistent with the subtle nature of the structural differences, their HPLC retention times using a slow gradient are very similar. Prior to heating, their UV-Vis spectra, which provide an accurate assessment of folding, are also virtually identical.

In contrast, following hydrogel formation, **2.1**-**2.3** can be easily distinguished on the basis of their hydrogel properties, CD spectra, and residual charge transfer absorbances. In other words, the aggregation that occurs upon heating amplifies tremendously the small structural differences between the molecules, a phenomenon that is best explained by proposing a highly ordered aggregate state conceptually analogous to protein derived amyloid formation. Taking a global view, our results indicate that

amyloid-like behavior, namely the existence of a stably folded state as well as the irreversible formation of a highly ordered aggregate involving an alternative conformation, is not unique to proteins, but may be a relatively general property of amphiphilic folding molecules in aqueous solution.

2.5 CONCLUSIONS

A series of structurally related amphiphilic foldamers were synthesized and their hydrogelation parameters were tested. Prior to heating, all foldamers of the series exhibited spectral characteristics consistent with folding in the pleated, stacked geometry characteristic of this class of foldamer. Following heating at 80° C, three of the four molecules exhibited irreversible aggregation to produce hydrogels. The hydrogels were characterized by rheology measurements and circular dichroism spectra revealed that hydrogel formation was dependent on highly ordered intermolecular assembly, conceptually analogous to protein amyloid formation. Hydrogel formation had the effect of amplifying the subtle structural differences between molecules, as the three amphiphilic foldamer constitutional isomers that formed hydrogels upon heating displayed significant differences in hydrogel properties. These results indicate that amyloid-like behavior is not unique to proteins, but may be a relatively general property of amphiphilic folding molecules in aqueous solution. Future work should include full characterization of the structure of the aggregated conformation on the atomic level by solid state NMR and fiber diffraction techniques.

2.6 EXPERIMENTAL METHODS

General Methods

Trifluoroacetic acid (TFA), piperidine, and N-methyl morpholine (NMM) were purchased from Sigma-Aldrich. All amino acid derivatives, benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate (PyBop), Hydroxybenzotriazole (HOBt) and Fmoc-Gly-Wang resin were purchased from Novabiochem. Compound **2.5** was obtained from Greg Gabriel of the Iverson group.

Monomer Synthesis

The Fmoc-Asp(tBu)-NDI monomer unit and Boc-DAN-OH was synthesized according to previously published methods. (Zych 2000) The Fmoc-R-DAN monomer units, where R is leucine, norleucine, isoleucine, or valine, were synthesized as described below.

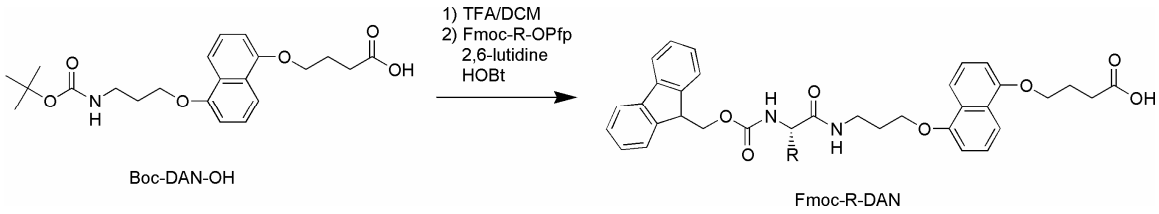


Figure 2.20 Synthesis of Fmoc-R-DAN monomer where R= Leu, Nle, and Ile.

Fmoc-Ile-Dan: Boc-DAN-OH (2.5 g, 5.4 mmol) was dissolved in a 1:1 mixture of DCM/TFA (8 mL) and stirred for 30 minutes at room temperature. The mixture was concentrated via rotary evaporation and azeotropically dried from n-heptane. The solid was dissolved in dry DMF (61 mL). To the solution, 2,6-lutidine (2.54 mL, 4 eq.) was added, followed by N-hydroxybenzotriazole (HOBt, 0.83 g, 1 eq.) and the pentafluorophenol ester of Fmoc-protected isoleucine (Fmoc-Ile-OPfp, 2.69 g, 0.95 eq.).

After stirring overnight at room temperature under a nitrogen atmosphere, the reaction mixture was partitioned between ethyl acetate and 0.2 M citric acid buffer, pH=4.5. The organic layer was extracted with buffer (3x), water (3x), and brine (1x), then dried over sodium sulfate. The solvent was evaporated under reduced pressure. The residue was precipitated with pentane and filtered. The filtrate was triturated with cold diethyl ether to produce a tan solid (3.26 g) in 94% yields. ^{1}H NMR δ (400 MHz, DMSO): 8.11 (t, 1H), 7.86 (d, 2H), 7.72 (m, 4H), 7.37 (m, 7H), 6.93 (d, 1H), 6.90 (d, 1H), 4.25 (m, 4H), 4.25 (t, 4H), 3.81 (t, 1H), 3.33 (m, 3H), 2.09 (m, 2H), 2.04 (m, 2H), 1.74 (m, 1H), 1.44 (m, 1H), 1.12 (m, 1H), 0.81 (m, 6H). ^{13}C NMR: 174.91, 172.07, 156.74, 154.58, 154.53, 144.50, 144.46, 141.41, 128.34, 127.73, 126.67, 126.10, 120.80, 114.30, 114.26, 106.36, 67.63, 66.30, 65.98, 47.39, 36.83, 36.15, 31.12, 29.52, 25.21, 25.00, 16.07, 11.48. HRMS (CI): calculated for $C_{38}H_{42}N_2O_7$ $(M+H)^{+1}$: m/z 639.3070, found 639.3064.

Fmoc-Leu-DAN: Synthesized according to the above procedure. ^{1}H NMR δ (400 MHz, DMSO): 8.06 (t, 1H), 7.86 (d, 2H), 7.71 (m, 4H), 7.48 (d, 1H), 7.33 (m, 6H), 6.93 (d, 1H), 6.90 (d, 1H), 4.25 (m, 4H), 4.11 (m, 4H), 4.01 (m, 1H), 3.35 (m, 2H), 2.07 (m, 2H), 1.97 (m, 2H), 1.50 (m, 4H), 0.84 (m, 6H). ^{13}C NMR: 174.91, 173.07, 156.60, 154.57, 154.54, 144.58, 144.43, 141.40, 128.30, 127.70, 126.67, 126.64, 126.07, 126.01, 125.99, 120.77, 120.75, 114.32, 114.26, 106.34, 67.61, 66.19, 65.99, 53.85, 47.38, 41.53, 36.29, 31.14, 25.01, 24.91, 23.62, 22.21. HRMS (CI): calculated for $C_{38}H_{42}N_2O_7$ $(M+H)^{+1}$: m/z 639.3070, found 639.3076.

Fmoc-Nle-DAN: Synthesized according to the above procedure. ^{1}H NMR δ (400 MHz, DMSO): 8.05 (t, 1H), 7.86 (d, 2H), 7.70 (m, 3H), 7.46 (d, 1H), 7.37 (m, 5H), 6.90 (m, 2H), 4.26 (m, 3H), 4.12 (m, 3H), 3.91 (q, 1H), 3.33 (m, 2H), 2.48 (m, 2H), 2.07 (m, 2H),

1.98 (m, 2H), 1.59 (bm, 1H), 1.50 (bm, 1H), 1.24 (m, 4H), 0.81 (m, 3H). ^{13}C NMR: 174.19, 172.03, 155.91, 153.88, 153.84, 143.87, 143.74, 140.68, 127.57, 126.98, 125.98, 125.95, 120.02, 113.63, 113.57, 105.60, 66.89, 65.55, 65.29, 64.88, 54.74, 46.67, 35.58, 31.70, 30.48, 28.88, 27.66, 24.32, 21.81, 13.82. HRMS (CI): calculated for $C_{38}H_{42}N_2O_7$ $(M+H)^{+1}$: m/z 639.3070, found 639.3074.

Fmoc-Val-DAN: Synthesized according to the above procedure. ^{1}H NMR δ (400 MHz, DMSO): 8.10 (t, 1H), 7.86 (d, 2H), 7.71 (m, 4H), 7.34 (m, 7H), 6.93 (d, 1H), 6.90 (d, 1H), 4.22 (m, 4H), 4.13 (t, 4H), 3.80 (t, 1H), 3.34 (m, 3H), 2.05 (m, 2H), 1.97 (m, 3H), 0.85 (m, 6H). ^{13}C NMR: 174.93, 171.98, 156.84, 154.58, 144.56, 141.38, 128.40, 128.34, 128.31, 127.74, 126.67, 126.64, 126.11, 126.08, 120.79, 114.33, 106.38, 67.75, 66.35, 65.99, 61.21, 47.38, 36.47, 31.16, 30.85, 29.60, 25.05, 19.96, 19.18. HRMS (CI): calculated for $C_{37}H_{41}N_2O_7$ $(M+H)^{+1}$: m/z 625.2914, found 625.2915.

Solid Phase Synthesis: All hexamers were synthesized using standard Fmoc solid phase peptide synthesis methods starting with Fmoc-Glycine substituted Wang resin and utilizing Fmoc-Asp(tBu)-NDI and Fmoc-R-DAN monomers with PyBop/HOBt/NMM as coupling reagents. The resin was capped with acetic anhydride after each coupling step. The product was cleaved using TFA/phenol as the cleavage cocktail. Crude product was purified via RP-HPLC (semi-preparative Vydac C18 peptide/protein column) using a linear gradient from 24% B to 46% B over 60 min, where solvent A is 10 mM ammonium acetate buffer and solvent B is acetonitrile. After lyophilization, peptide was desalted using a Waters SepPak C18 cartridge and eluted with 1:1 H_2O/acetonitrile.

Compound **2.1**: Ac(Leu-DAN-Asp-NDI)$_3$Gly-OH (ESI) calculated for $C_{142}H_{149}N_{19}O_{39}$ $(M-2H)^{-2}$: m/z 1372.014, found 1372.008.

Compound **2.2**: Ac(Nle-DAN-Asp-NDI)$_3$Gly-OH (ESI) calculated for $C_{142}H_{149}N_{19}O_{39}$ $(M-2H)^{-2}$: m/z 1372.014, found 1372.017.

Compound **2.3**: Ac(Ile-DAN-Asp-NDI)$_3$Gly-OH (ESI) calculated for $C_{142}H_{148}N_{19}O_{39}$ $(M-3H)^{-3}$: m/z 914.340, found 914.338.

Compound **2.4**: Ac(Val-DAN-Asp-NDI)$_3$Gly-OH (ESI) calculated for $C_{139}H_{143}N_{19}O_{39}$ $(M-2H)^{-2}$: m/z 1350.990, found 1350.991.

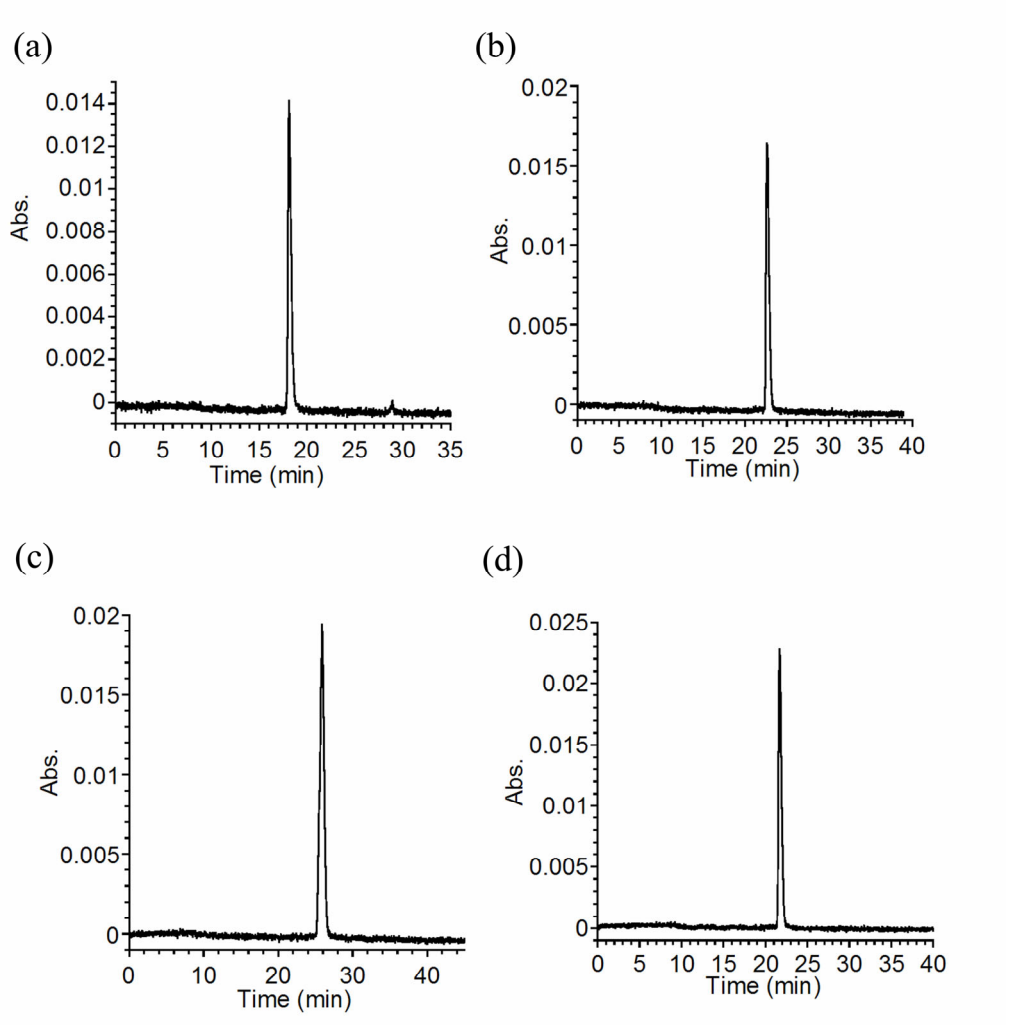


Figure 2.20 HPLC chromatograms of purified compounds **2.1**-**2.4** (a-d).

HPLC

HPLC was performed on a Waters HPLC system equipped with a photodiode detector and Vydac C18 peptide/protein analytical reversed phase column with a gradient of 22% B to 44% B over 60 minutes, where solvent A is 10 mM ammonium acetate buffer and solvent B is acetonitrile. All chromatograms were monitored at 384 nm.

UV Spectroscopy

Ultraviolet spectra (UV) were recorded on a Perkin-Elmer Lambda 35 spectrophotometer for the series of aedamers **2.1**-**2.4** (15 μM) in 50 mM Phosphate buffer at pH 7.0 in a 1 cm cuvette. Cetyl trimethylammonium bromide (CTAB) was added as a solid to achieve a 2% (w/v) solution and the UV spectrum was recorded. Observed percent hypochromism is reported as [(Abs_{382} with CTAB – Abs_{386} without CTAB)/ Abs_{382} with CTAB] x 100.

Light Scattering

Light scattering was performed on a Wyatt Technologies DynaPro MS dynamic light scattering instrument. Measurements were obtained using filtered, 2.5 mM samples in 50 mM sodium phosphate buffer at pH 7.0 and 100 mM sodium chloride in a microcuvette. Data was analyzed assuming a globular protein structure with cumulants analysis utilizing software provided by the manufacturer.

Visible Spectroscopy

Visible spectra were recorded on a Perkin-Elmer Lambda 35 spectrophotometer. Solutions of compounds **2.1-2.4** at 2.5 mM concentration were dissolved in 50 mM sodium phosphate buffer at pH 7.0 with 100 mM NaCl, filtered using a 0.45 μm nylon

filter, and placed in a 1 mm pathlength cell. An initial spectrum was recorded and the samples were heated in an 80˚ C water bath for 1.5 hours and cooled to room temperature before a final spectrum was taken.

Rheology

Dynamic frequency, time, temperature and strain sweep experiments were performed on an Anton Paar Physica MCR 300 rheometer equipped with a Peltier temperature controller and a 25 mm diameter parallel plate. All experiments were performed at 25˚ C except the temperature sweep experiment was performed during a ramp of 25˚ C to 80˚ C. Gels were formed in the rheometer by placing a sample of 4 mM preformed gel in the rheometer, lowering the top plate, and using capillary action to fill the 1 mm gap with a 4 mM solution of compound. The plate was covered with silicon oil and the rheometer was heated to 80˚ C for 1 hour. Gel strength was characterized by the storage, G', and loss, G", moduli via frequency sweep experiments at a constant strain of 0.5%. Shear thinning studies were performed by measuring the viscosity while changing the shear rate from 0.1 to 100 s^{-1}. Recovery of the elastic properties of the gel was measured by applying a 1000% strain for 180 s at a frequency of 6 Hz then reducing the strain to 0.5% at the same frequency and monitoring the storage and loss moduli for 15 minutes.

Circular Dichroism

Circular dichroic experiments were performed on a Jasco J-815 circular spectropolarimeter and collected at 25˚ C in a 1 mm quartz cell. Samples were measured at 0.2 mM concentration in 50 mM sodium phosphate buffer at pH 7.0 with 100 mM NaCl added. After an initial spectrum was recorded, the samples were heated to 80˚ C for

1 hour using the Peltier temperature controller. The samples were then cooled to 25˚ C before the final spectrum was recorded.

Time trial experiments for compounds **2.1-2.3** were performed by heating 0.2 mM aedamer with 10% 0.2 mM pre-gelled material added in 50 mM sodium phosphate, 100 mM NaCl buffer at pH 7.0 in a sealed vial at 25° C and 37° C for 2 weeks. The samples were cooled to 25° C before the final spectrum was recorded in a 1 mm quartz cell.

SEM

Hydrogels were applied to copper grids by dipping the grids in preformed gel samples. The coated grids were dipped in deionized water to desalt the gel and frozen by submerging in liquid nitrogen. The grids were lyophilized to dryness and coated with 8 nm Pt/Pd. Images were obtained with a Zeiss Supra 40 VP scanning electron microscope at 5 kV using either an SE2 or in-lens camera.

CHAPTER 3

Aromatic Donor-Acceptor Interaction in Peptide Folding

3.1 CHAPTER SUMMARY

Introduction

Nature utilizes a variety of non-covalent interactions to modulate protein folding, including aromatic interactions. Although homo-aromatic interactions are often found with phenylalanine and tryptophan, there are far less hetero-aromatic interactions in natural proteins, and due to the lack of an electron-deficient aromatic amino acids, there are no aromatic donor-acceptor interactions. Due to the strengthened electrostatic component, it is possible that a donor-acceptor type of aromatic interaction may lend itself to more stable protein structures, higher selectivity, and higher affinity of protein-protein and protein-ligand interactions. This chapter describes the design and synthesis of non-natural amino acids with NDI- and DAN-functionalized side chains for use in a β-hairpin secondary structural model.

Goals

The hydrophobic driving force and electrostatic complementarity of the DAN-NDI interaction has often been used to form new abiotic folding topologies in water. (Lokey 1995, Gabriel 2005) In addition, this same principle lends itself for potential use in peptide and protein folding. Experiments in this chapter were designed to pursue the question: Can strategic placement of aromatic donor and acceptor units stabilize protein or peptide folded structure? The short term goal of this research was to design a model

system for a secondary structural motif and synthesize non-natural amino acids for incorporation of DAN and NDI units into peptides.

Approach

A model β-hairpin peptide developed by Marcey Waters was chosen for incorporation of non-natural NDI and DAN containing amino acids to determine if aromatic donor-acceptor interactions could stabilize a secondary structural polypeptide motif. Molecular modeling was performed to determine the molecular requirements of the non-natural amino acid side chains for optimal face-centered donor-acceptor stacking. Two sets of NDI and DAN non-natural amino acids were designed and synthesized for use in the model system.

Results

Overall, the work in this chapter describes the efforts to synthesize the amino acids required for use in solid phase peptide synthesis of the designed model β-hairpin peptide. While the N_α-Boc protected NDI and DAN amino acids could be synthesized and used for Boc solid phase peptide synthesis, the transformation of the N_α-Cbz derivative into N_α-Fmoc dipeptides for Fmoc solid phase peptide synthesis was thwarted by inherent insolubility or organogel behavior in organic solvents.

3.2 BACKGROUND

3.2.1 α-Helix and β-Hairpin Secondary Structural Models

It is now widely accepted that aromatic-aromatic interactions contribute to the overall stability of a variety of protein tertiary and quaternary structures. However, the

magnitude of those contributions differs between structures due to differences in packing, solvation, and sequence of the individual proteins studied. In order to avoid the complications that arise from these differences, aromatic-aromatic interactions have been studied in smaller polypeptides that adopt well-defined secondary structural motifs including α-helices and β-hairpins.

In α-helices, the role of aromatic interactions have been studied by identifying interacting pairs of phenylalanine residues in α-aminoisobutyryl (Aib)-containing peptides. The Aib residues at specific locations within the sequence provides conformational rigidity, resulting in helical structures. The crystal structure of the peptides, which placed the phenylalanine (Phe) residue at the *i* and *i+4* positions and the *i* and *i+2* positions within the helix, showed a variety of conformations of aromatic interactions. When placed in the *i* and *i+4* positions, the phenylalanine residues were displayed on the same side of the helix and showed intramolecular parallel-displaced, edge-to-face, and inclined arrangements. In the *i* and *i+2* positions, the phenylalanine residues resided on opposite faces of the helix and displayed intermolecular edge-to-face interactions only. (Aravinda 2003)

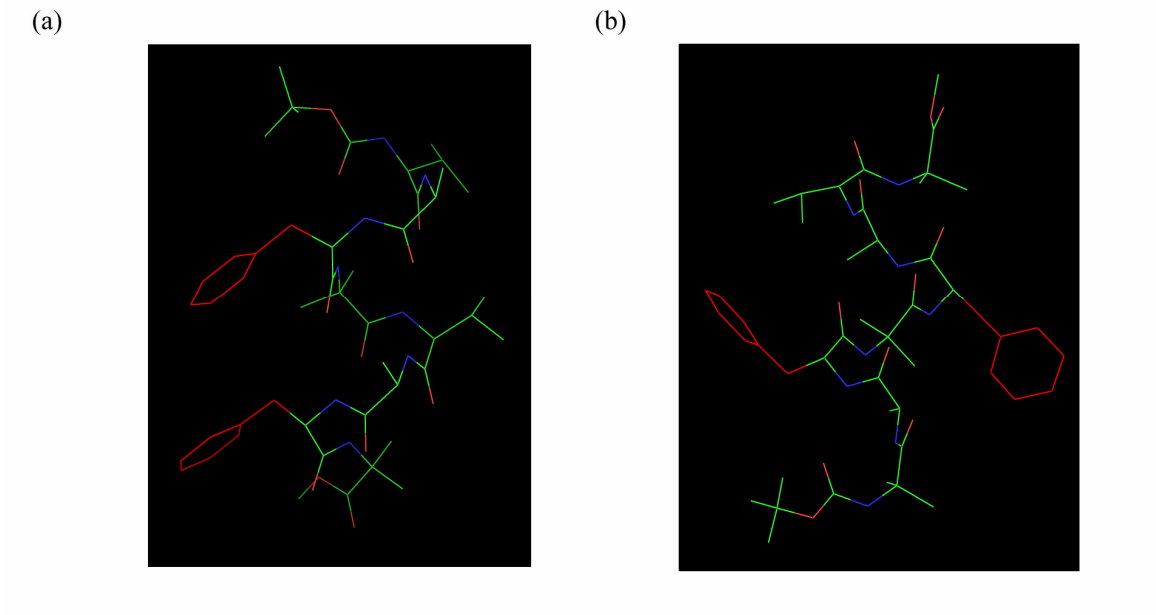


Figure 3.1 Aib-containing α-helix displaying phenylalanine in the (a) *i* and *i+4* positions, with an intramolecular inclined arrangement and (b) *i* and *i+2* positions for intermolecular aromatic interactions. (adapted from Aravinda 2003)

While the crystal structures in Figure 3.1 provided information about the conformations adopted by aromatic pairs in α-helices, the thermodynamic contribution of aromatic interactions in α-helices has been measured by determination of the helical content of poly-alanine peptides containing phenylalanine at the *i* and *i+4* positions. In addition to phenylalanine pairs at the *C*-terminus and in the middle of the helix sequence, pentafluorophenylalanine was also tested as a pair with phenylalanine. Interestingly, these studies showed that the phenylalanine pair and the phenylalanine-pentafluorophenylalanine pair stabilized the helix to a similar degree (ΔG = -0.27 kcal/mol) even though the later maximized the complimentary quadrupole interaction through a face-to-face conformation. Through modeling, however, the face-to-face conformation was determined to be inaccessible in the *i* and *i+4* positions of the α-helix

whereas offset-stacked and edge-to-face geometries were accessible, thereby explaining the smaller than expected contribution of the electrostatically matched pair. (Butterfield 2002)

In β-hairpins, the crystal structures of several *de novo* designed β-hairpins have shown cross-strand aromatic interactions between phenylalanine residues. (Aravinda 2002, Karle 2000) The ability of phenylalanine residues to interact in a β-hairpin was dependent on the position of the residues. In a peptide where there was a close approach of phenylalanine residues placed directly across the strand at non-hydrogen bonding positions in the β-hairpin, the centroid to centroid distances were between 5.37 and 5.28 Å and the interplanar angle was between 47 and 63°, suggesting an inclined arrangement. (Figure 3.2a) In a peptide where the phenylalanine residues were placed in a diagonal orientation across the strand, the centroid distance were 6.12 Å with an interplanar angle of 86.98°, which was an edge-to-face conformation similar to the majority of Phe-Phe interactions in natural proteins. (Aravinda 2004)

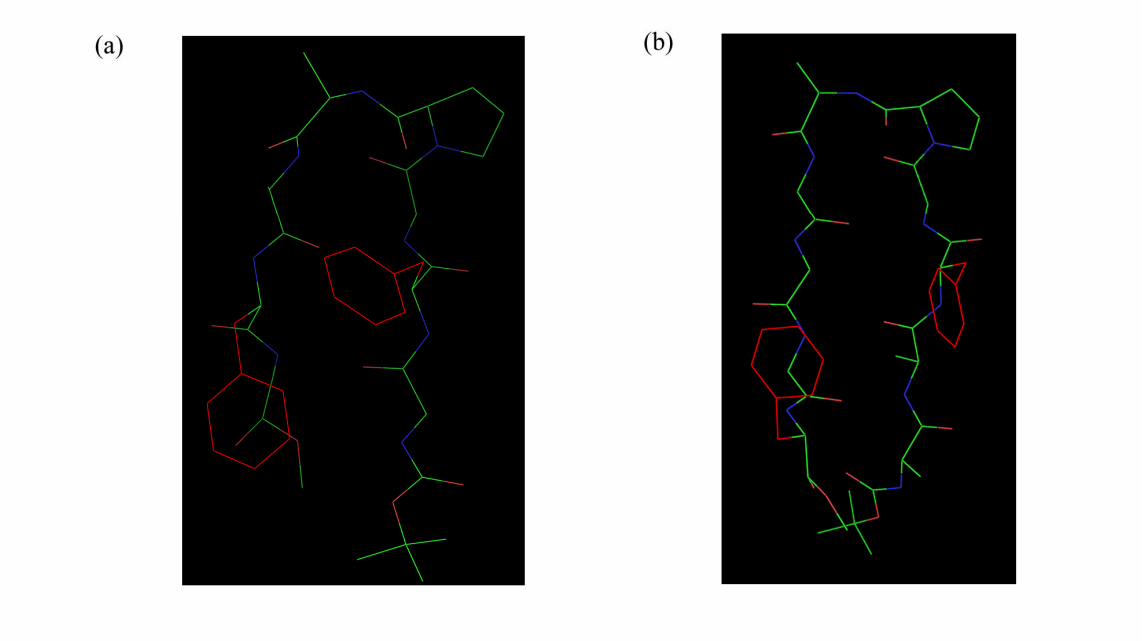


Figure 3.2 Crystal structure of de novo designed β-hairpin displaying (a) close approach and (b) diagonal orientation of a phenylalaine pair. (Cambridge Crystalographic Data Centre CCDC 218048, CCDC 218050) (Aravinda 2004)

One of the more stable small β-hairpins known to date is the 16 amino acid trpzip peptide (Figure 3.3). The trpzip peptide was based on residues 41-56 of the B1 IgG-binding domain of protein G (gb1 peptide). The β-hairpin fold of gb1 peptide is stabilized by a cluster of four hydrophobic residues: tryptophan, tyrosine, phenylalanine, and valine. By replacing the phenylalanine, tyrosine, and valine with tryptophan residues, the hairpin character increased greatly, the thermal denaturing curve became more cooperative and the melting temperature of the tryptophan substituted peptide was higher by at least 40° C. The NMR structure of the trpzip peptide displayed edge-to-face packing of the tryptophan residues. (Cochran 2001)

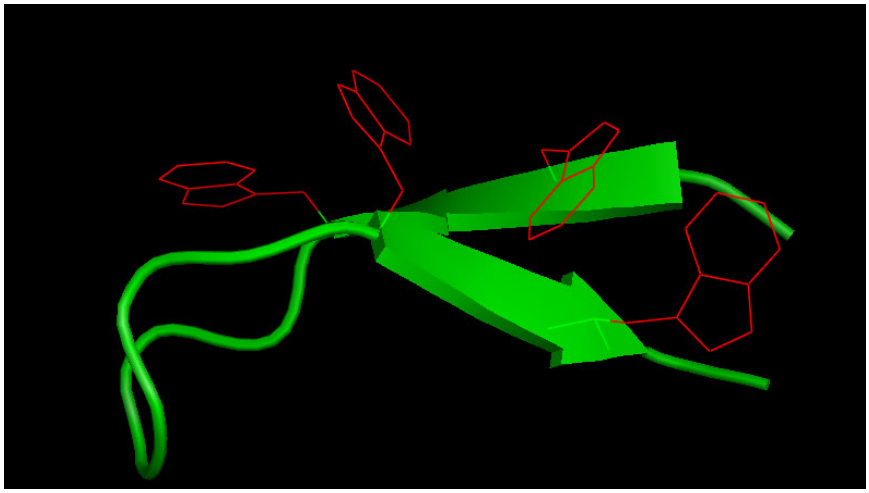


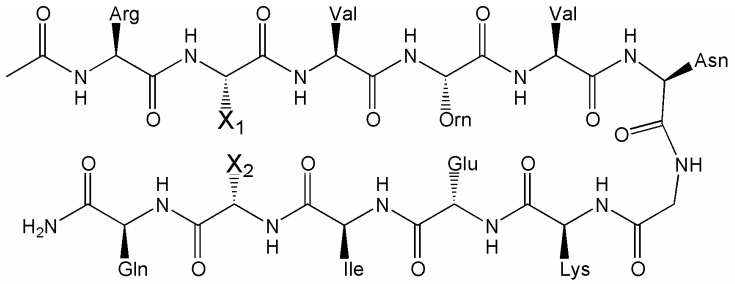
Figure 3.3 NMR structure of the TrpZip4 peptide displaying a cluster of tryptophan residues with edge-to-face packing. (PDB 1LE3, Cochran 2001)

In a body of work focused on a designed β-hairpin, Waters has measured the effect of a variety of aromatic interactions on β-hairpin folding. (Waters 2004) The β-hairpin used for all the studies involved an Asn-Gly turn sequence, which has a propensity to form hairpins via a Type I' turn. (Stanger 1998) A type I' turn is defined by the phi/psi angles of the two center residues of a 4-residue turn sequence: $\phi_{i+1} = +60°$, $\varphi_{i+1} = +30°$, $\phi_{i+2} = +90°$, $\varphi_{i+2} = 0°$. For solubility and avoidance of aggregation, the peptide was overall +2 charged. The percent folded of the peptides was determined by the differences in the chemical shifts of the H_α resonances and the glycine H_α resonances in the turn region as compared to a fully folded control, and a fully unfolded control. In addition, thermodynamic analyses were determined by the temperature dependence of the chemical shifts. While the sequence of the β-hairpin was similar for the different interactions measured, the optimal positions of the interacting residues differed depending on the specific interaction.

The first interaction to be measured was the Phe-Phe interaction. The phenylalanine residues were positioned in the close approach directly across the strand as

shown in Figure 3.4a. The Phe-Phe interaction was compared with the purely hydrophobic cyclohexylalanine (Cha)-cyclohexylalanine interaction. The experiments showed a preference of -0.55 kcal/mol for self-association of the Phe-Phe interaction over the Phe-Cha interaction. In addition, NOEs suggested an edge-to-face conformation of the phenylalanine rings. (Tatko 2002)

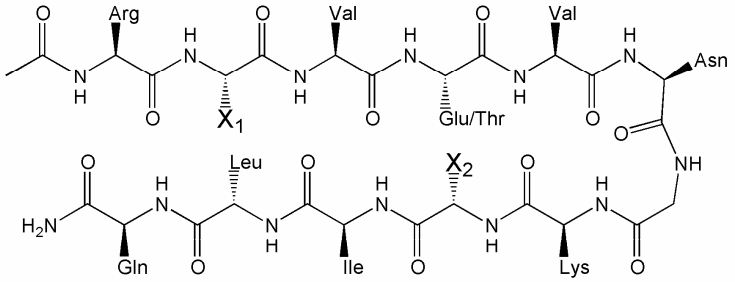
(a)



2,11-substituted: Ac-Arg-**X_1**-Val-Orn-Val-Asn-Gly-Lys-Glu-Ile-**X_2**-Gln-NH_2

1 ⟶ **12**

(b)



2,9-substituted: Ac-Arg-**X_1**-Val-Thr/Glu-Val-Asn-Gly-Lys-X_2-Ile-Leu-Gln-NH_2

1 ⟶ **12**

Figure 3.4 Waters' β-hairpin sequences with (a) 2,11- substitution and (b) 2,9-substitution of aromatic interactions.

In addition to the Phe-Phe aromatic-aromatic interaction, cation-π, C-H-π, and N-H-π interactions were measured with the β-hairpin system. The cation-π interaction was determined between tryptophan and lysine or methylated lysine. (Tatko 2004, Hughes 2007) The most stable hairpin utilized the interaction between tryptophan and tri-

methylated lysine and increased the overall stability by -0.7 kcal/mol over the lysine-containing hairpin. (Hughes 2005) The C-H-π interaction was measured with glucosylated serine (Ser(Glc)) and tryptophan. The magnitude of the interaction was measured to be -0.8 kcal/mol greater than unglycosylated serine, which was greater than the Phe-Phe or cation-π interactions described above. (Kiehna 2007) The N-H-π interaction was measured between tryptophan and acylated lysine and was determined to be similar to the interaction of tryptophan and lysine even with the loss of a positive charge (-0.34 versus -0.3 kcal/mol, respectively). (Hughes 2006)

In order to accommodate the additional size of the side chains used for these studies, the β-hairpin was altered such that the residues of the interaction being measured were diagonal across the strand instead of directly across as shown in Figure 3.4b. In order to retain an overall charge of +2, the residue in position 4 was chosen as glutamic acid if the interaction being measured involved a positively charged residue.

3.2.2 Non-Natural Aromatic Donor-Acceptor Amino Acids

While the interaction of aromatic donor-acceptor amino acids have not been measured directly in peptide systems, non-natural amino acids with aromatic acceptor side chains have appeared in the literature with several applications. Non-natural NDI amino acids have been used in a self-assembled cyclic peptide for NDI excimer formation (Figure 3.5a). The efficiency of charge transfer of the self-assembled structure was dependent on the length of the NDI side chain and the number of NDI units in the structure. (Horne 2005)

Non-natural NDI amino acids have also been used to design and synthesize synthetic pores with sticky π-clamps in rigid-rod β-barrels. The tetrameric pores in the "open" position allowed translocation of molecules through the pores. Modeling suggested that the NDI clamps can "close" to catch analytes as they pass through the pore

(Figure 5b). π-Clamping was shown to occur with a variety of analytes such as nucleotides, electron-rich and electron-deficient naphthalenes. Nucleotide clamping showed discrimination for purines over pyrimidines with excellent sensitivity. (Tanaka 2007) This work was expanded for sensing applications including artificial tongues. Using the pores to mediate the release of 5(6)-carboxyfluorescein from self-quenched vesicles in combination with reactive amplification of analytes from a variety of food, the pores were used to detect levels of sucrose in Coca-Cola, lactose in milk, acetate in rice vinegar, lactate in spoiled milk, citrate in orange juice, and glutamate in soy sauce. In several cases such as sucrose, no amplification was required as the analyte binds directly with sufficient affinity and selectivity. In cases such as glutamate, a reactive amplifier was required. The glutamate analyte was converted to α-ketoglutarate by a transaminase which reacted with a DAN hydrizide. The DAN was then recognized by the NDI units in the synthetic pores with higher selectivity and affinity. (Litvinchuk 2007) The DAN-α-ketoglutarate conjugate is shown bound to the synthetic pore by molecular modeling in Figure 3.5c.

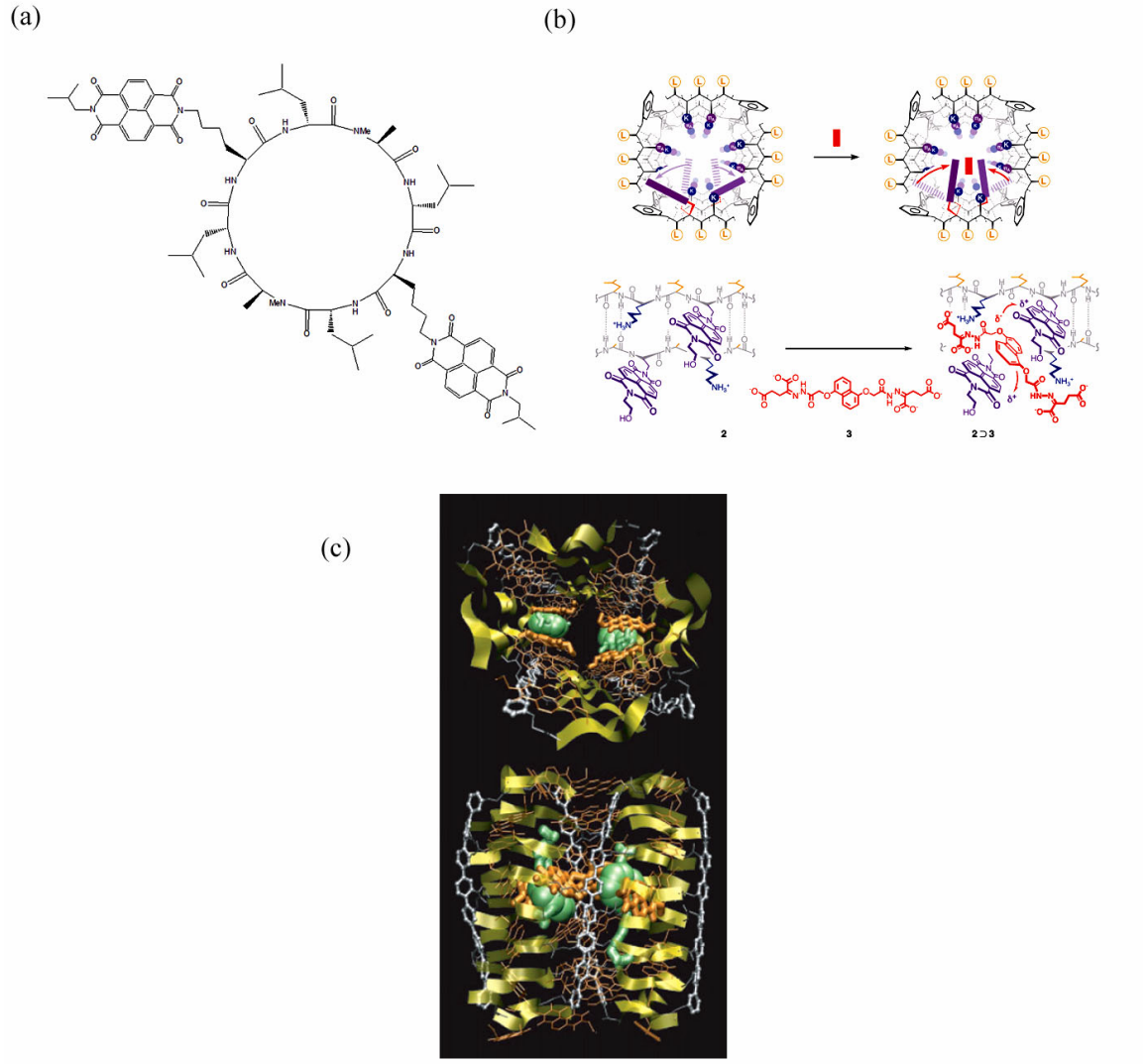


Figure 3.5 (a) Cyclic peptide utilizing NDI amino acids for charge transfer in self-assembled cylinders. (b) Synthetic pores containing NDI amino acids for capture of a DAN guest. (c) Molecular modeling with a top and side view of a synthetic pore binding a reactive amplified ketoglutarate for detection of glutamate in soy sauce. (Reproduced with permission from Nature Publishing Group) (Litvinchuk 2007)

While aromatic-aromatic interactions occur often in protein folding, donor-acceptor interactions do not exist in natural proteins due to the lack of naturally occurring

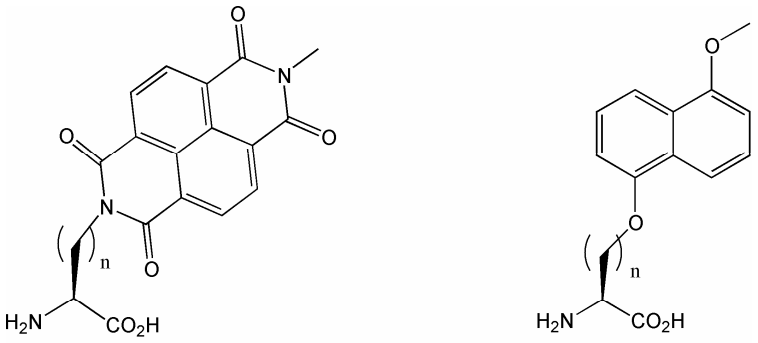
electron-deficient aromatic amino acids. Due to the strengthened electrostatic component, it is possible that aromatic donor-acceptor interactions may provide a larger interaction than naturally occurring aromatic-aromatic interactions. In addition, previous work has shown that monomers of NDI and DAN units have an interaction of -3.1 kcal/mol. (Gabriel 2002) This is comparable to the energy of a hydrogen bond, which is generally thought to be from 1 to 5 kcal/mol in proteins. By designing and synthesizing non-natural amino acids containing NDI and DAN side chains, the ability of an aromatic donor-acceptor interaction to stabilize protein structure can be measured by incorporation into a model system.

3.3 RESULTS

3.3.1 β-Hairpin and Amino Acid Design

The model system chosen for determining the ability of NDI and DAN containing amino acids to stabilize protein and peptide structures was Waters' β-hairpin design in Figure 3.4. From a qualitative comparison of the interactions measured and the size of the NDI and DAN molecules, it was unclear which orientation, directly across the strand in positions 2 and 11 (figure 3.4a) or diagonal across the strand in positions 2 and 9 (figure 3.4b), would provide the best accessibility for the optimal face-centered stacking conformation of the NDI and DAN moieties. In addition, the synthetic feasibility of the NDI and DAN-containing amino acids needed to be considered. The amino acids chosen for incorporation into the β-hairpin peptide are shown below in Figure 3.6. With the amino acid design chosen, molecular modeling was then performed on both the 2,9-substituted and 2,11-substituted hairpin using amino acids **3.1** and **3.3** as shown in Figure 3.7. While the 2,11-substituted hairpin showed promise for face-to-face stacking, the optimal geometry, where the side chains of the NDI and DAN were almost orthogonal in

space, was not accessible. Modeling of the 2,9-substituted hairpin using **3.1** and **3.3** showed optimal geometry of the NDI and DAN relative to each other, however, the side chains of **3.1** and **3.3** were too short. Utilization of amino acids **3.2** and **3.4** in the 2,9-substituted hairpin showed the optimal face-centered stacking as shown in Figure 3.7d.



n = 1 **3.1**
n = 2 **3.2**

n = 1 **3.3**
n = 2 **3.4**

Figure 3.6 Design of NDI and DAN non-natural amino acids for incorporation in a β-hairpin model system.

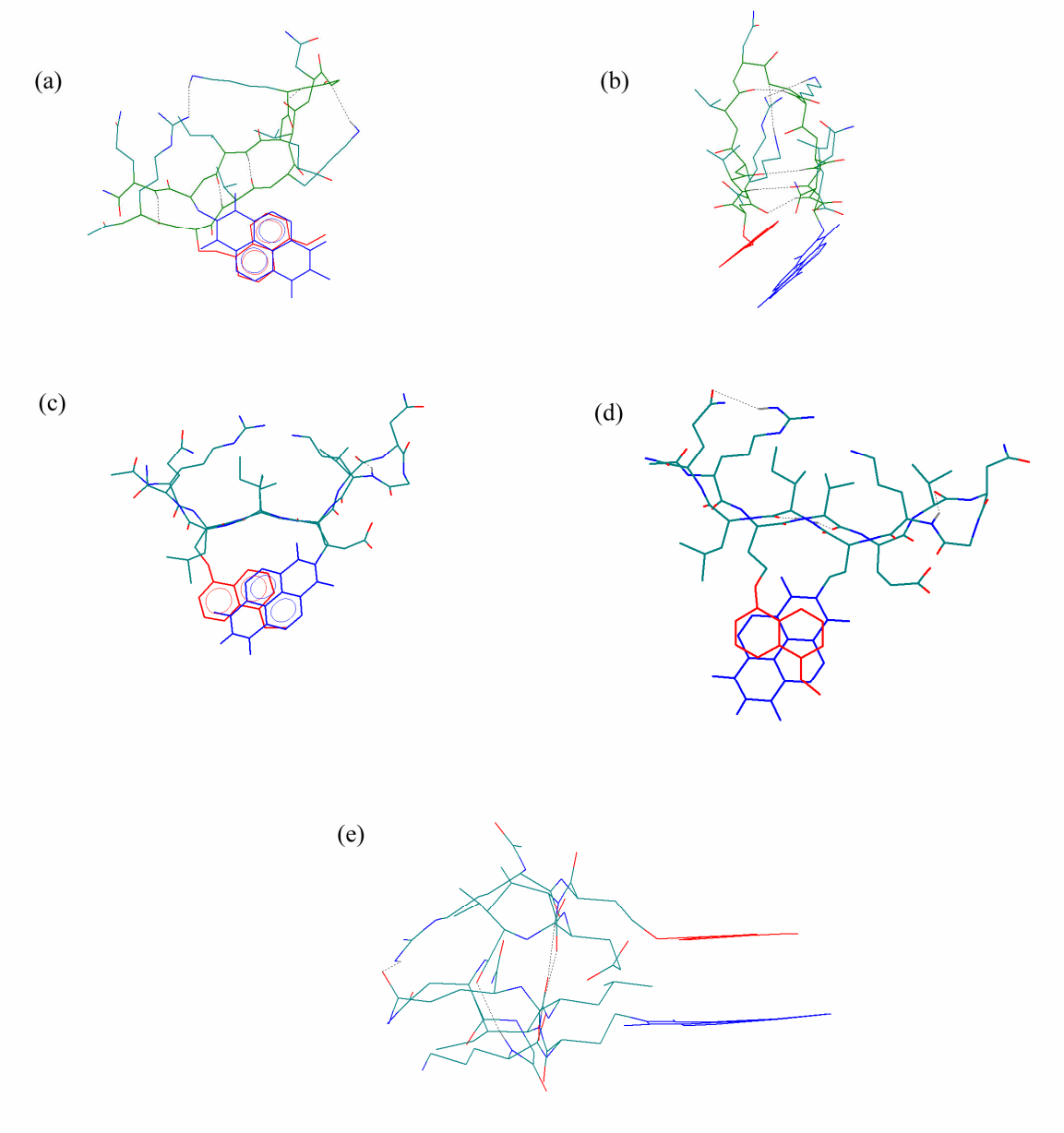


Figure 3.7 Molecular modeling of the β-hairpin system. (a) 2,11-substituted hairpin with **3.1** and **3.3** displaying suboptimal orientation. (b) View of 2,11-substituted hairpin with **3.1** and **3.3** displaying aromatic overlap. (c) 2,9-substituted hairpin with **3.1** and **3.3**. (d) 2,9-substituted hairpin with **3.2** and **3.4** displaying optimal orientation. (e) 2.9-substituted hairpin with **3.2** and **3.4** displaying optimal aromatic overlap.

As a starting point, the sequences for the 2,9-substituted hairpin with amino acids **3.2** and **3.4** were designed as shown in Figure 3.8. In order to determine the fraction of the β-hairpin that is folded, four peptides were designed. Peptide **3.5** is the peptide of interest, while peptide **3.6** is the fully folded peptide. The peptide is constrained to a fully folded population using a disulfide bridge between two terminal cysteine residues. Peptides **3.7** and **3.8** represents the fully unfolded peptide. By synthesizing the two strands of the hairpin individually, the chemical shifts of the α protons in an unfolded state can be measured without concern of partially folded hairpins. The fraction folded can be determined from the following equation:

$$\text{Fraction folded} = f = [\delta_{obs}-\delta_0]/[\ \delta_{100}-\delta_0]$$

The fraction of the population in the folded hairpin can be used to determine the ΔG of the interaction using the following equations:

$$K = f/(1-f)$$

$$\Delta G = -RT \ln K$$

Thermodynamic analysis can be performed using the temperature-dependence of the glycine chemical shift or the average Hα chemical shift for Val3, Val5, Lys8, and Ile10 in peptide **3.5** conjunction with the following equations:

$$\Delta\delta = \Delta\delta_{max}[e^{(x/RT)}]/[1+e^{(x/RT)}]$$

$$x=T(\Delta S°_{298}+\Delta Cp°\ln(T/298))-(\Delta H°_{298}+\Delta Cp°(T-298))$$

3.5 Ac-R**X**^{1}VTVNGK**X**^{2}ILQ-NH_2

3.6 Ac-**C**R**X**^{1}VTVNGK**X**^{2}ILQ**C**-NH_2 X^1=DAN, X^2=NDI

3.7 Ac-R**X**^{1}VTVNG-NH2

3.8 Ac-NGK**X**^{2}ILQ-NH2

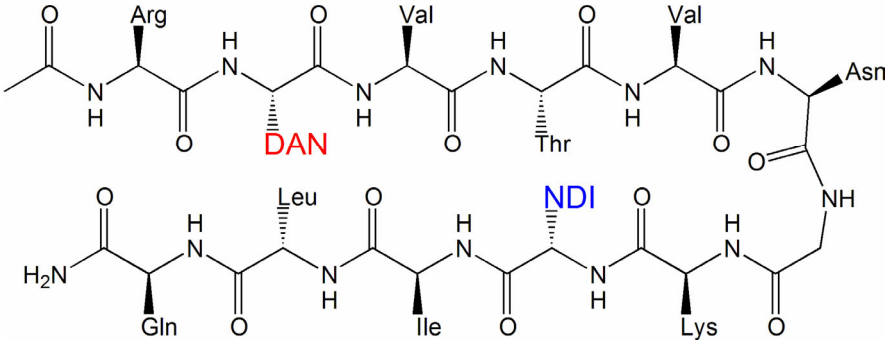
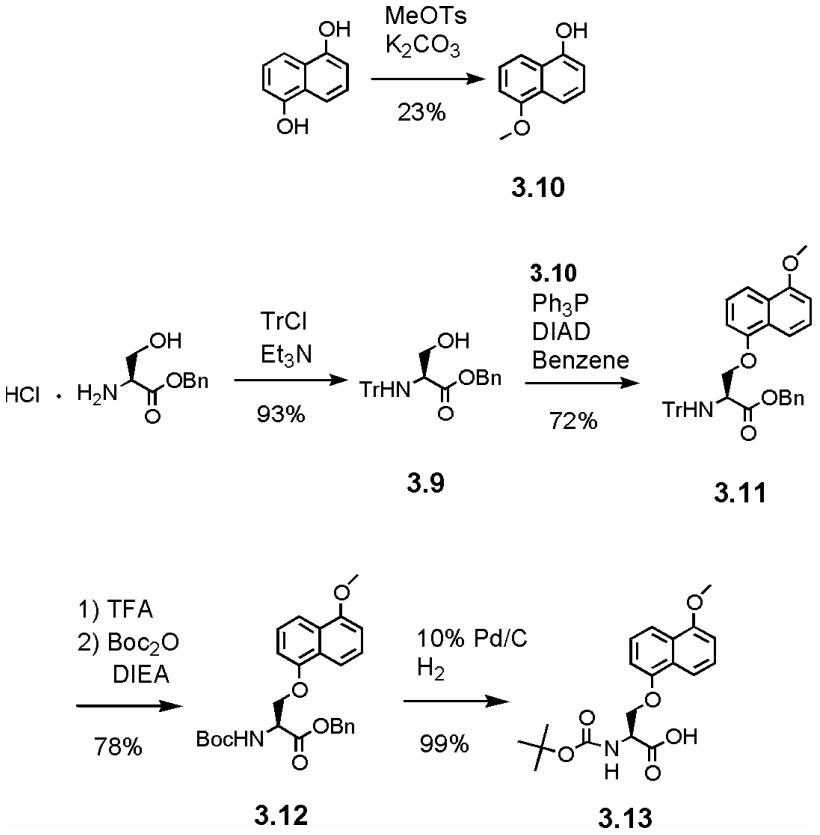


Figure 3.8 Peptide sequences to analyze the fraction folded and the thermodynamic parameters of the DAN-NDI interaction.

3.3.2 Amino Acid Synthesis

Serine benzyl ester was the starting material for both the DAN and NDI amino acids, **3.1** and **3.3**, and protected as the *N*-trityl derivative as shown in Scheme 3.1. In addition, 1,5-dihydroxynaphthalene was mono-alkylated with methyl *p*-toluenesulfonate to give compound **3.10**. The key transformation in the synthesis of **3.1** was the Mitsunobu reaction of mono-alkylated dihydroxynaphthalene **3.10** with trityl-serine benzyl ester **3.9**. Serine derivatives substituted with a β-leaving group primarily undergo β-elimination even with strong nucleophiles under S_N2. The Mitsunobu reaction has been shown to proceed in excellent yields with strong nuclophiles; however, it proceeds in poor yields with weak nucleophiles unless serine is *N*-protected with a trityl or phenylluorenyl group. It is known that *N*-trityl and *N*-phenylfluorenyl group protect the C_α of amino acids from base-promoted racemization, presumably by sterically hindering deprotonation of the α-hydrogen. (Cherney 1996) The bulkiness of the trityl protecting group was required for

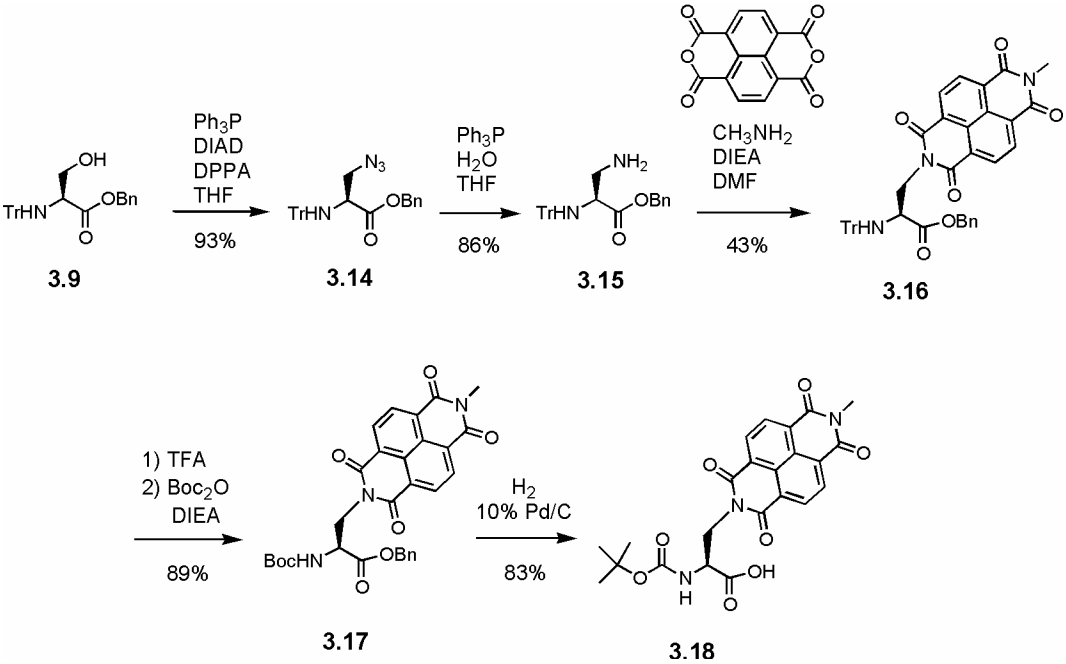
successful Mitsunobu coupling in this case, as the corresponding Boc derivative gave only the dehydroalanine product due to β-elimination. After deprotection of the trityl group with TFA and Boc protection, the benzyl ester was removed using hydrogenation conditions to afford the Boc-DAN amino acid **3.13**.



Scheme 3.1 Synthesis of Boc protected DAN amino acid **3.13** for the 2,11-substituted hairpin.

The key intermediate in the synthesis of the NDI amino acid was the formation of a protected diaminoproprionic acid (Dap) amino acid. This was accomplished in high yield from the trityl protected serine benzyl ester **3.9** via Mitsunobu coupling with diphenylphosphoryl azide as a stable azide source (Scheme 3.2). The azidoalanine **3.14** was then converted to the trityl protected Dap benzyl ester **3.15** by a Staudinger

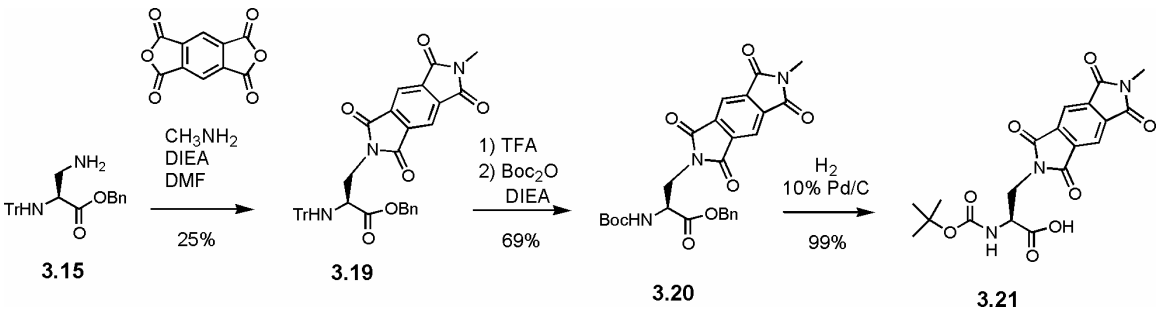
reduction. The most difficult synthetic step was formation of the asymmetric diimide **3.16**. The conventional synthesis of the asymmetric diimide involved refluxing two primary amines with 1,4,5,8-naphthalenetetracarboxylic dianhydride. This method, however, produced particularly low yields of no more than 35%. Recently, a microwave-assisted synthesis of asymmetrical naphthalenediimides was discovered. (Pengo 2006) A modified version of this procedure improved the yield of **3.16** to 45%. After trityl group deprotection and Boc protection, the benzyl ester was removed by hydrogenation to afford the Boc-NDI amino acid **3.18**.



Scheme 3.2 Synthesis of Boc protected NDI amino acid **3.18** for the 2,11-substituted hairpin.

In addition to Boc protected NDI amino acid **3.18**, a non-natural amino acid with a smaller electron-deficient side chain was synthesized. As mentioned in Chapter 1, the electron-deficient aromatic unit PDI can be used in place of NDI for donor-acceptor interactions, however, the PDI is smaller and does not match the surface area of the DAN unit as well as NDI. The smaller hydrophobic surface area of the PDI unit allows it to be

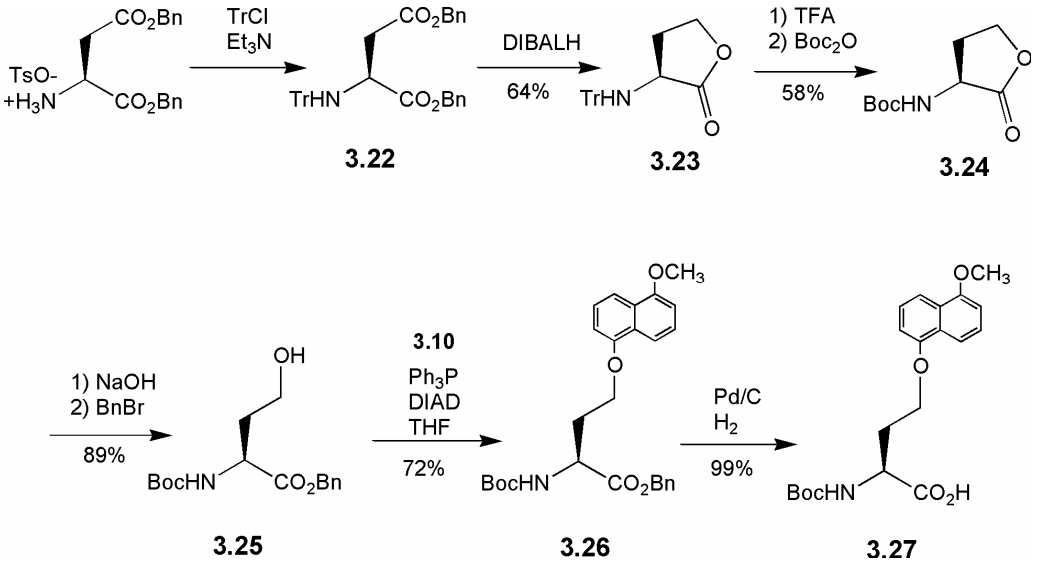
less prone to aggregation from self-stacking. The trityl protected Dap benzyl ester **3.15** from synthesis of the NDI amino acid **3.18** was used with 1,2,4,5-benzenetetracarboxylic dianhydride and methylamine to produce the trityl protected PDI benzyl ester **3.19** by conventional heating in DMF. After deprotection of the trityl group followed by reprotection with the Boc protecting group, the Boc protected PDI benzyl ester **3.20** was subjected to hydrogenation conditions to remove the benzyl ester which produced the Boc protected PDI amino acid **3.21**.



Scheme 3.3 Synthesis of the alternate electron-deficient amino acid, Boc protected PDI **3.21**.

The ability to synthesize the DAN and NDI amino acids **3.2** and **3.4** in a fashion analogous to **3.1** and **3.3** was prevented by the extreme expense of the common amino acid starting material, homoserine. The key intermediate in the synthesis of DAN amino acid **3.2** was Boc-homoserine benzyl ester **3.25**. Aspartic acid dibenzyl ester was trityl protected to afford derivative **3.22**. (Scheme 3.4) Using DIBALH, the diester was reduced and cyclized to the γ-lactone **3.23** in moderate yield. (Baldwin 1987) While the trityl protected lactone underwent ring opening with sodium hydroxide, alkylation of the carboxylate salt with benzyl bromide was thwarted by reformation of the lactone. Conversion of the trityl protecting group to the corresponding Boc derivative allowed conversion of the lactone to the Boc-homoserine benzyl ester **3.25** in good yield.

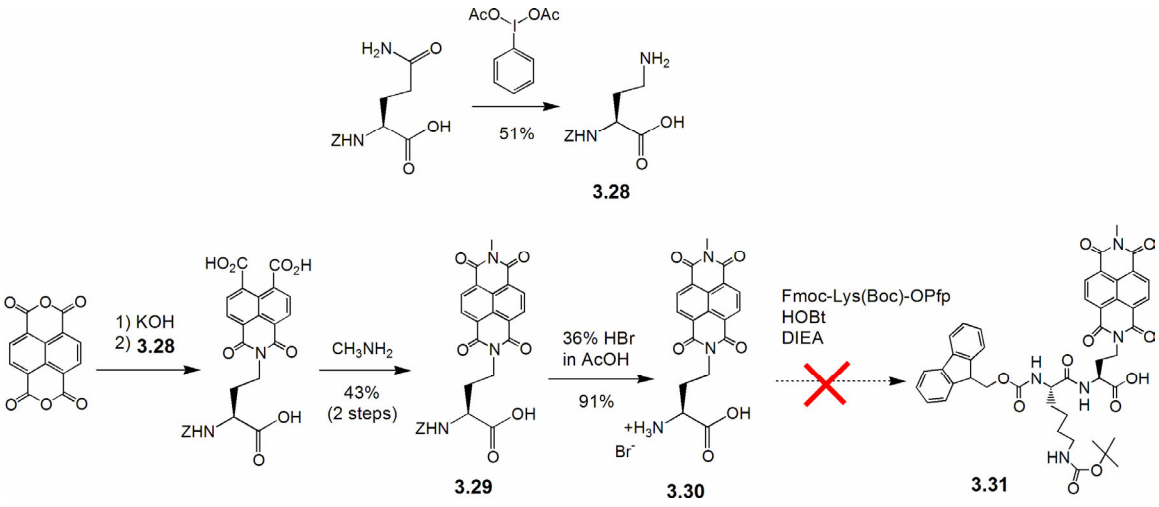
(Ozinskas 1986) It is possible that the carbamate provided hydrogen bonding partners for the resulting alcohol to prevent reformation of the lactone. While the trityl protection was nessesary to prevent β-elimination of compound **3.9**, the additional methylene of homoserine eliminated the necessity for the trityl protection for the Mitsunobu coupling with compound **3.25**. Hydrogenation of the benzyl ester provided the Boc protected DAN amino acid **3.27**.



Scheme 3.4 Synthesis of Boc protected DAN amino acid **3.27** for the 2,9-substituted hairpin.

While construction of NDI amino acid **3.3** relied on the key intermediate diaminoproprionic acid, construction of **3.4** relied on the key intermediate diaminobutyric acid (Dab) **3.28**. Formation of Cbz-protected Dab **3.28** was achieved by a Hoffmann rearrangement of Cbz-glutamine with iodosobenzene diacetate as shown in Scheme 3.5. (Zhang 1997, Macedo 2002) Refluxing of the Cbz-Dab free acid with 1,4,5,8-naphthalenetetracarboxylic dianhydride and methylamine produced dismal yields (17%) of the NDI product **3.29**. Alternatively, formation of the mono-anhydride in aqueous

base, followed by refluxing at pH 6.4 with **3.28** afforded the monoimide by precipitation at more acidic pH. The monoimide was then converted to diimide **3.29** by heating with methylamine in DMF in an overall yield of 43%. (Horne 2005) In order to incorporate the NDI amino acid into the β-hairpin by Fmoc solid phase peptide synthesis, the free amino acid **3.30** would have to be coupled to the pentafluorophenol ester of the *N*-terminally adjacent Fmoc amino acid, lysine (Fmoc-Lys(Boc)-OPfp). While soluble at high temperature and during column purification, Cbz derivative **3.29** was highly insoluble in organic solvents and water when pure. It was discovered that compound **3.29** was only soluble in a dilute 10-20% methanol in dicholormethane mixture and upon standing for several hours, formed an organogel. The gelling behavior was clear from the NMR, which even taken immediately after dissolution, showed very broad peaks with no distinct splitting patterns. Although **3.29** was insoluble, hydrogenation of the Cbz group was attempted to no avail. As an alternate deprotection method, a 36% solution of HBr in glacial acetic acid was added to **3.29** and the suspension was stirred for 1 hour to afford the deprotected compound **3.30**. (Ben-Ishai 1952) While the HBr salt of compound **3.30** was perfectly soluble in organic solvents, the addition of base to provide the free amine for amide coupling gave completely insoluble material in most organic solvents and organogels in a few others depending on the concentration. Even heating with Fmoc-Lys(Boc)-OPfp at 80° C overnight gave little coupling product, all of which was Fmoc deprotected.



Scheme 3.5 Synthesis of Fmoc-Lys(Boc)-NDI-OH dipeptide **3.31** for the 2,9-substituted hairpin.

3.4 DISCUSSION

3.4.1 β-Hairpin Design

While the β-hairpin design had been successfully used to measure a variety of aromatic interactions, there was some concern whether its use would allow accurate measurement of the interaction between NDI and DAN side chains as it can be sensitive to the orientation of the NDI and DAN units in the stacking conformation. In addition, the NDI and DAN moieties have a larger surface area than most of the aromatic amino acids. This could potentially drive aggregation of the hairpins to form a self-assembled material such as a fiber or hydrogel. While interesting, measurement of the interaction may not be possible at concentrations required for NMR.

From the molecular modeling, it was clear that the system with the optimal orientation of the NDI and DAN side chains was the 2,9-substituted hairpin with the amino acids **3.2** and **3.4** diagonally placed across the hairpin. Not only did the model show optimal overlap of the aromatic units, the model displayed optimal orientation of

the aromatic units with almost orthogonal display of the side chains. The 2,11-substituted hairpin with the amino acids **3.1** and **3.3** directly across the strand was not completely ruled out by modeling, however. The modeling showed optimal overlap of the aromatic units, but not optimal orientation. As to not rely solely on computer modeling, the 2,11-substituted hairpin should still be synthesized and may provide additional information on how orientation of the units effects the magnitude of the interaction.

3.4.2 Amino Acid Synthesis

The major roadblock in the synthesis of the amino acids required for incorporation into the β-hairpin was the organogel behavior and insolubility of the NDI free amino acid **3.30**. It is expected that the free amino acid of the shorter side chain derivative **3.18** would display similar behavior. Organogel behavior was not completely unexpected in this case. As discussed in Chapter 1, self-stacked NDI units cover most of the hydrophobic aromatic face through an electrostatically favorable offset face-to-face conformation. In addition, it is feasible that the free base, carboxylate of **3.30** can be involved with hydrogen bonding. It is not far-fetched to propose that the combination of hydrogen bonding and the NDI self-stacking may have led to a well-defined self-assembled material in organic solvents.

Interestingly, compound **3.29** has been used previously in the synthetic pores discussed earlier. While not specifically mentioned, a similar problem may have occurred. Instead of Fmoc solid phase peptide synthesis, the authors performed a solution phase synthesis where a *C*-terminal tripeptide carboxyamide was coupled to compound **3.29** followed by hydrogenation of the Cbz protecting group and an addition solution phase coupling to an *N*-terminal leucine. (Tanaka 2007) A similar, but modified strategy would be needed to make the β-hairpin. Due to the length of the β-hairpin peptide, a full solution phase strategy would not be feasible. However, compound **3.29** could be coupled

as the preformed pentafluorophenol ester to the *C*-terminally adjacent residue isoleucine. After hydrogenation, the dipeptide would be coupled to Fmoc-Lys(Boc)-OPfp to form the tripeptide that could then be used in Fmoc solid phase peptide synthesis. Alternately, it has been shown that Cbz protected amino acids can be used in Fmoc solid phase synthesis with the selective removal of the Cbz group with diethylaluminum chloride without resin cleavage. (Yoo 2007) This would allow compound **3.29** to be used directly in solid phase synthesis, possibly with the use of microwave assisted synthesis. Additionally, the synthesis could be retooled so that the starting material, glutamine, is Boc protected instead of Cbz protected. The Boc group may provide better solubility as seen with compound **3.18** and would also provide the complete set of Boc protected amino acids **3.13**, **3.18**, **3.21**, **3.27** and **3.29** for the use of Boc solid phase peptide synthesis. While the PDI amino acid **3.21** was synthesized for comparison with the extent of donor-acceptor interaction between the NDI and DAN units, the small surface area may avoid the solubility and organogelation issues of the NDI unit and prove to be a suitable alternative.

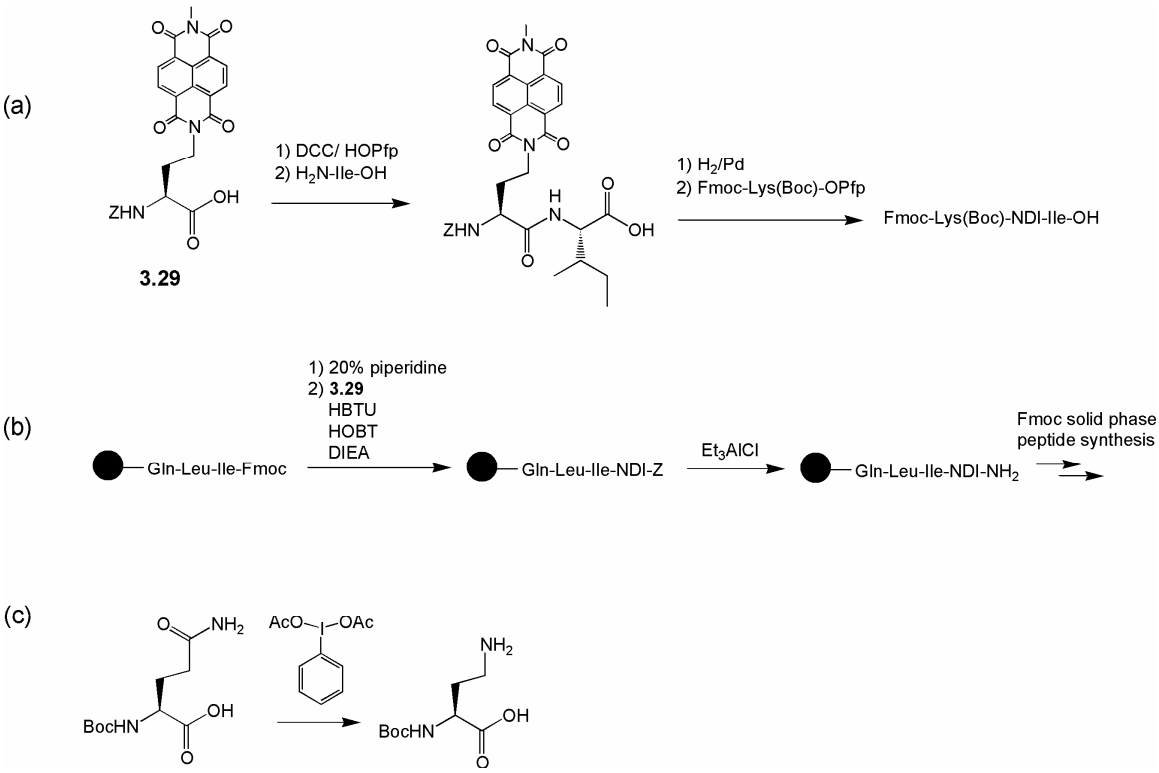


Figure 3.9 Alternative synthetic routes to incorporation of NDI amino acid **3.4** in the 2,9-substituted hairpin.

3.5 CONCLUSIONS

β-Hairpins have been used as a model system to measure a wide variety of aromatic interactions in peptide and protein folding. Results of previous work suggest that aromatic donor-acceptor interactions could be a valuable addition to the list of aromatic interactions that provide protein stability. Modeling has shown that the 2,9-subustituted hairpin could provide optimal orientation and overlap of the aromatic faces of the NDI and DAN side chains through placement of the amino acids in positions diagonal across the hairpin strands.

The Fmoc-protected dipeptide containing a DAN side chain for incorporation into the 2,9-substituted hairpin has been synthesized. In addition, Boc-protected NDI and

DAN containing amino acids with a side chain one methylene shorter have also been synthesized for incorporation into the 2,11-substituted hairpin. Unfortunately, an Fmoc protected dipeptide containing an NDI side chain for the 2,9-substituted hairpin could not be synthesized due to insolubility and organogel formation. This behavior is a testament to the strength of the aromatic-aromatic interactions of the NDI self-association and presents a strong argument for the potential success of using NDI-DAN interactions to stabilize peptide folding. Alternate methods to synthesize acceptable derivatives have been proposed including coupling the C-terminally adjacent residue prior to Cbz deprotection followed by tripeptide formation, use of the Cbz protected amino acid directly in Fmoc solid phase peptide synthesis, or using Boc-glutamine as a starting material for the synthesis to afford a residue compatible with Boc solid phase peptide synthesis.

3.6 EXPERIMENTAL METHODS

General Procedures All starting materials were purchased from Aldrich and Novabiochem and used without further purification. Flash chromatography was performed on Silicycle Silia-P silica gel. Microwave reactions were performed in a CEM Mars microwave synthesis module equipped with a temperature probe. ^{1}H and ^{13}C NMR were recorded on a Varian Unity Plus 400 MHz spectrophotometer in the indicated solvent. Chemical shifts are expressed in parts per million (ppm, δ) relative to tetramethylsilane (TMS) (δ = 0.00 ppm).

Molecular modeling Molecular modeling was performed in Hyperchem 7 (Hypercube Inc., Gainesville, FL) with the Amber force field. The hairpin was geometry optimized

separately from the NDI and DAN side chains while constraining the turn region to standard phi/psi angles for a type I' turn. The NDI and DAN side chains were then geometry optimized in the 2,9- and 2,11-substituted hairpins followed by addition geometry optimization of the entire hairpin structure.

N-Trityl serine benzyl ester (3.9) Serine benzyl ester hydrochloride (5 g, 0.021 mol) was suspended in dry DCM (45 ml) under nitrogen gas. Triethylamine (6.4 ml, 0.043 mol) was added dropwise to the suspension and the mixture was cooled to 0° C in an ice bath. Trityl chloride (6.3 g, 0.022 mol) was dissolved in dry DCM (13 ml) and added to the reaction mixture dropwise. The reaction was stirred 6 hours at 0° C under nitrogen gas. The solvent was removed via rotary evaporation. The residue was partitioned between ethyl acetate and saturated sodium bicarbonate solution. The organic layer was extracted with saturated $NaHCO_3$ (3 x 100 ml), 0.1 M citric acid (100 ml), and water (100 ml) then dried over Na_2SO_4. The ethyl acetate was removed under reduced pressure. Purification of the crude product by column chromatography (9:1 hexanes/EtOAc) provided compound **3.9** as a white foam (8.7 g, 93%). ^{1}H NMR ($CDCl_3$) δ 7.51 (m, 5H), 7.35-7.30 (m, 15H), 4.77 (d, *J* = 12 Hz, 1H), 4.66 (d, *J* = 12 Hz, 1H), 3.74 (m, 1 H), 3.59 (m, 2 H), 3.06 (bs, 1H), 2.36 (t, *J* = 8 Hz, 1H). ^{13}C NMR ($CDCl_3$) δ 173.59, 145.86, 135.55, 128.99, 128.76, 128.52, 128.32, 128.21, 126.89, 71.29, 67.13, 65.18, 58.09. HRMS (CI) calculated for $C_{29}H_{29}NO_3$ $(M+H)^{+1}$: m/z 438.2069, found 438.2074.

5-Methoxynaphthalene-1-ol (3.10) 1,5-Dihydroxynaphthalene (10 g, 62 mmol) was dissolved in a suspension of K_2CO_3 (19.4 g, 140 mmol) in acetonitrile (155 ml). Methyl p-toluenesulfonate (4.7 ml, 31 mmol) was added to the suspension and the reaction was refluxed for 5 hours under nitrogen gas. The mixture was concentrated under reduced

pressure and the residue was dissolved in DCM (200 ml) and filtered through Celite. After concentration by rotary evaporation, the residue was dissolved in 5% aqueous NaOH (200 ml) and filtered. The filtrate was acidified to pH 2 with concentrated HCl and extracted with DCM (3 x 100 ml). The combined organic layers were concentrated under reduced pressure. Purification of the crude product by column chromatography (CHCl3) provided compound **3.10** as a grey solid (1.19 g, 23%). 1H NMR (CDCl3) δ 7.87 (d, *J* = 8 Hz, 1H), 7.76 (d, *J* = 8 Hz, 1H), 7.40 (t, *J* = 8 Hz 1H), 7.30 (t, *J* = 8 Hz, 1H), 6.86 (d, *J* = 8 Hz, 2H), 5.33 (bs, 1H), 4.00 (s, 3H). 13C NMR ($CDCl_3$) δ 155.33, 151.11, 126.90, 125.26, 125.11, 114.68, 113.59, 109.41, 104.41, 55.53. HRMS (CI) calculated for $C_{11}H_{11}O_2$ $(M+H)^{+1}$: m/z 175.0759, found 175.0762.

(S)-Benzyl 3-(5-methoxynaphthalen-1-yloxy)-2-(tritylamino)propanoate (3.11) Compound **3.9** (4.1 g, 9.4 mmol), triphenylphosphine (2.7 g, 10.3 mmol), and compound **3.10** (2.4 g, 14 mmol) were dissolved in dry benzene (107 ml). While the mixture was stirring under nitrogen atmosphere, diisopropylazodicarboxylate (2 ml, 10.3 mmol) was added dropwise. The reaction was stirred overnight under nitrogen. The mixture was concentrated under reduced pressure. Purification of the crude product by column chromatography (95:5 hexanes/EtOAc) gave compound **3.11** as a white solid (4.1 g, 72%). ^{1}H NMR ($CDCl_3$) δ 7.92 (d, *J* = 8 Hz, 1H), 7.76 (d, *J* = 8 Hz, 1H), 7.67-7.65 (m, 5H), 7.36-7.21 (m, 16 H), 6.87 (t, *J* = 8 Hz, 2H), 4.78 (d, *J* = 12 Hz, 1H), 4.61 (d, *J* = 12 Hz, 1H), 4.53 (q, *J* = 4 Hz, 1H), 4.32 (t, *J* = 8 Hz, 1H), 4.02 (s, 3H), 3.15 (bs, 1H). ^{13}C NMR ($CDCl_3$) δ 173.71, 155.01, 153.82, 145.71, 135.14, 128.71, 128.29, 128.09, 128.03, 127.93, 126.53, 126.48, 126.45, 125.18, 124.90, 114.58, 114.37, 105.26, 104.47, 70.98, 70.49, 66.97, 56.19, 55.42. HRMS (CI) calculated for $C_{40}H_{35}NO_4$ $(M+H)^{+1}$: m/z 593.2566, found 593.2568.

(S)-Benzyl 2-(*tert*-butoxycarbonyamino)-4-(5-methoxynaphthalen-1-yloxy) propanoate (3.12) Compound **3.11** (0.12 g, 0.17 mmol) was dissolved in DCM (2.5 ml). Methanol (37 µl) and trifluoroacetic acid (38 µl, 0.5 mmol) were added to the solution and the reaction was stirred for 10 minutes. The reaction was concentrated under reduced pressure. The residue was dissolved in DCM (1.6 ml) with diisopropylethylamine (86 µl, 0.5 mmol) added. Di-tert-butyl dicarbonate (0.04 g, 0.18 mmol) was added to the mixture and the reaction was stirred overnight under nitrogen gas. The reaction mixture was washed with 0.5 M citric acid (3 x 10 ml), saturated $NaHCO_3$ (3 x 10 ml), and water (10 ml). The organic layer was concentrated under reduced pressure. Purification of the crude product by column chromatography (70:25:5 DCM/hexanes/ACN) gave compound **3.12** as a clear oil (0.061 g, 67%).^{1}H NMR ($CDCl_3$) δ 7.93 (d, *J* = 8 Hz, 1H), 7.74 (d, *J* = 8 Hz, 1H), 7.47 (m, 2H), 7.27-2.19 (m, 5H), 6.86 (d, *J* = 8 Hz, 1H), 6.79 (d, *J* = 8 Hz, 1H), 5.78 (d, *J* = 8.4 Hz, 1H), 5.26 (q, *J* = 13.6 Hz, 2H), 4.90 (m, 1H), 4.55 (d, *J* = 6.4 Hz, 1H), 4.42 (d, *J* = 6.4 Hz, 1H), 3.99 (s, 3H), 1.51 (s, 9H). ^{13}C NMR ($CDCl_3$) δ 169.96, 155.25, 155.01, 153.43, 135.00, 128.31, 128.10, 127.87, 126.43, 126.23, 125.34, 124.79, 114.93, 113.89, 105.45, 104.47, 80.10, 68.60, 67.28, 55.33, 53.71, 28.17. HRMS (CI) calculated for $C_{26}H_{30}NO_6$ $(M+H)^{+1}$: m/z 452.2073, found 452.2070.

(S)-2-(*tert*-butoxycarbonylamino)-3-(5-methoxynaphthalen-1-yloxy)propanoic acid (3.13) Compound **3.12** (0.19 g, 0.4 mmol) was dissolved in methanol (4 ml). 10% Palladium on carbon (22 mg, 5 mol%) was added and the solution was purged with hydrogen gas. The reaction was stirred for 5 hours under a balloon of hydrogen. The mixture was filtered and rinsed with methanol. The combined filtrates were concentrated

under reduced pressure. The residue was dissolved in a small volume of DCM and precipitated with hexanes. Filtration of the precipitate gave compound **3.13** as a white solid (0.14 g, 99%). ^{1}H NMR ($CDCl_3$) δ 8.86 (bs, 1H), 7.82 (d, *J* = 8.4 Hz, 1H), 7.66 (d, *J* = 8.4 Hz, 1H), 7.30 (t, *J* = 8 Hz, 1H), 7.21 (m, 1H), 6.72 (m, 2H) 5.63 (d, *J* = 4 Hz, 1H), 4.82 (m, 1H), 4.48 (d, *J* = 6.4 Hz, 1H), 4.35 (d, *J* = 6.4 Hz, 1H), 3.94 (s, 3H), 1.44 (s, 9H). ^{13}C NMR ($CDCl_3$) δ 174.59, 155.59, 155.03, 153.49, 126.49, 126.29, 125.41, 124.85, 115.05, 113.85, 105.59, 104.55, 80.64, 68.43, 55.44, 53.52, 28.12. HRMS (CI) calculated for $C_{19}H_{24}NO_6$ $(M+H)^{+1}$: m/z 362.1604, found 362.1599.

(S)-Benzyl 3-azido-2-(triylamino)propanoate (3.14) Compound **3.9** (8.9 g, 20.4 mmol) was dissolved in dry THF (204 ml). Triphenylphosphine (5.6 g, 21.4 mmol) was added and the mixture was cooled to 0° C in an ice bath. Diisopropylazodicarboxylate (4.2 ml, 21.4 mmol) was added dropwise followed by dropwise addition of diphenylphosphorylazide (4.8 ml, 21.4 mmol). The reaction was allowed to warm to room temperature and stirred overnight under nitrogen. The reaction mixture was concentrated under reduced pressure. Purification of the crude product by column chromatography (95:5 hexanes/EtOAc) gave compound **3.14** as a yellow oil (8.7 g, 93%). ^{1}H NMR ($CDCl_3$) δ 7.49 (m, 5H), 7.28 (m, 15H), 4.72 (d, *J* = 12 Hz, 1H), 4.58 (d, *J* = 12 Hz, 1H), 3.62 (m, 1H), 3.34 (m, 2H), 2.95 (d, *J* = 9.7, 1H). ^{13}C NMR ($CDCl_3$) δ 172.28, 145.39, 143.52, 135.75, 135.02, 129.25, 128.61, 128.56, 128.51, 128.44, 128.38, 128.32, 128.27, 128.24, 128.16, 128.11, 127.95, 127.58, 126.85, 126.60, 74.32, 71.23, 67.07, 66.60, 56.54, 54.91, 31.71, 28.76. HRMS (CI) calculated for $C_{29}H_{27}N_4O_2$ $(M+H)^{+1}$: m/z 463.2134, found 463.2131.

(S)-Benzyl 3-amino-2-(triylamino)propanoate (3.15) Compound **3.14** (7.8 g, 16.8 mmol) was dissolved in THF (48 ml). Triphenylphosphine (5.3 g, 20.2 mmol) was added slowly and the solution was stirred at room temperature until generation of nitrogen gas ceased. The reaction was refluxed for 2 hours. Water (7.8 ml) was added to the reaction and the reaction was refluxed overnight. The mixture was concentrated under reduced pressure. The crude mixture was purified by column chromatography (95:5 DCM/triethylamine) and concentrated. The solid was triturated with diethyl ether and filtered twice to remove the remaining triphenylphosphine oxide. The pure compound **3.15** was concentrated under reduced pressure to give a clear oil (6.3 g, 86%). ^{1}H NMR ($CDCl_3$) δ 7.50 (m, 5H), 7.27-7.15 (m, 15H), 4.68 (d, *J* = 12 Hz, 1H), 4.51 (d, *J* = 12 Hz, 1H), 3.71 (m, 1H), 2.91 (m, 2H). ^{13}C NMR ($CDCl_3$) δ 174.17, 146.05, 135.65, 133.19, 132.37, 132.27, 132.20, 132.17, 129.10, 129.06, 128.99, 128.94, 128.91, 128.80, 128.73, 128.68, 128.63, 128.52, 128.48, 128.42, 128.39, 128.34, 128.29, 128.27, 128.24, 128.23, 128.20, 128.10, 128.05, 127.79, 127.19, 126.75, 71.25, 66.82, 58.57, 46.99. HRMS (CI) calculated for $C_{29}H_{29}N_2O_2$ $(M+H)^{+1}$: m/z 437.2230, found 437.2239.

(S)-benzyl 3-(7-methyl-1,3,6,8-tetraoxo-7,8-dihydrobenzo[*lmn*][3,8]phenanthrolin-2(1*H*,3*H*,6*H*)-yl)-2-(tritylamino)propanoate (3.16) Compound **3.15** (0.9 g, 2.0 mmol) and 1,4,5,8-naphthalenetetracarboxylic acid anhydride (0.54 g, 2 mmol) were dissolved in DMF (20 ml). Diispropylethylamine (0.36 ml, 2 mmol) was added and the mixture was subjected to microwave heating to 140° C for 10 minutes. After cooling to room temperature, 2.0 M methylamine in THF (2 ml, 4 mmol) was added and the reaction was stirred capped for 5 hours. The mixture was then heated to 140° C for 1 hour. Purification of the crude product by column chromatography (70:25:5 DCM/hexanes/ACN) gave compound **3.16** as a yellow solid (0.61 g, 43%). ^{1}H NMR ($CDCl_3$) δ 8.68 (d, *J* = 7.6 Hz,

2H), 8.62 (d, *J* = 7.6 Hz, 2H), 7.52-7.49 (m, 5H), 7.22-7.12 (m, 10H), 6.99-6.82 (m, 5H), 4.60 (m, 2H), 4.52 (d, *J* = 12.4 Hz, 1H), 4.26 (d, *J* = 12.4 Hz, 1H), 3.92 (m, 1H), 3.61 (s, 3H), 3.00 (d, *J* = 10.8 Hz, 1H). ^{13}C NMR ($CDCl_3$) δ 172.79, 163.23, 162.91, 145.79, 135.00, 131.23, 131.05, 128.98, 128.76, 128.11, 128.07, 127.95, 127.19, 126.76, 126.64, 126.50, 71.33, 67.21, 54.81, 44.03, 27.68. HRMS (CI) calculated for $C_{44}H_{34}N_3O_6$ $(M+H)^{+1}$: m/z 700.2448, found 700.2446.

(S)-benzyl 2-(*tert*-butoxycarbonylamino)-3-(7-methyl-1,3,6,8-tetraoxo-7,8-dihydrobenzo[*lmn*][3,8]phenanthrolin-2(1*H*,3*H*,6*H*)-yl)propanoate (3.17) Compound **3.16** (0.18 g, 0.25 mmol) was dissolved in DCM (4 ml). Methanol (80 μl) and trifluoroacetic acid (60 μl, 0.75 mmol) were added. The reaction was stirred 10 minutes. The reaction was concentrated under reduced pressure. The residue was dissolved in DCM (3 ml) with diispropylethylamine (135 μl, 0.28 mmol) added. Di-tert-butyl dicarbonate (62 mg, 0.28 mmol) was added to the mixture and the reaction was stirred overnight under nitrogen gas. The reaction mixture was washed with 0.5 M citric acid (3 x 10 ml), saturated $NaHCO_3$ (3 x 10 ml), and water (10 ml). The organic layer was concentrated under reduced pressure. Purification of the crude product by column chromatography (70:25:5 DCM/hexanes/ACN) gave compound **3.17** as a white solid (0.128 g, 89%). ^1H NMR ($CDCl_3$) δ 8.68 (s, 4H), 7.26 (m, 2H), 7.19 (m, 3H), 5.42 (d, *J* = 8.4 Hz, 1H), 5.16 (s, 2H), 4.91 (m, 1H), 4.55 (d, *J* = 5.6 Hz, 2H), 3.56 (s, 3H), 1.18 (s, 9H). ^{13}C NMR ($CDCl_3$) δ 170.05, 162.92, 162.87, 155.21, 134.85, 131.06, 130.81, 128.35, 128.25, 128.16, 126.58, 126.42, 126.32, 126.17, 79.95, 67.63, 51.81, 41.78, 41.78, 27.95, 27.36. HRMS (CI) calculated for $C_{30}H_{28}N_3O_8$ $(M+H)^{+1}$: m/z 558.1876, found 558.1874.

(S)-2-(*tert*-butoxycarbonylamino)-3-(7-methyl-1,3,6,8-tetraoxo-7,8-dihydrobenzo[*lmn*][3,8]phenanthrolin-2(1*H*,3*H*,6*H*)-yl)propanoic acid (3.18) Compound **3.17** (0.24 g, 0.4 mmol) was dissolved in DMF (4 ml). 10% Palladium on carbon (44 mg, 10 mol%) was added and the solution was purged with hydrogen gas. The reaction was stirred under a balloon of hydrogen for 5 hours at room temperature. The solution was filtered and the solid rinsed with DMF. The filtrate was concentrated under reduced pressure to afford compound **3.18** as a grey solid (0.16 g, 83%). ^{1}H NMR (DMSO) δ 8.53 (s, 4H), 7.05 (d, *J* = 8.4Hz, 1H), 4.36 (m, 3H), 3.37 (s, 3H), 1.12 (s, 9H). HRMS (CI) calculated for $C_{23}H_{22}N_3O_8$ $(M+H)^{+1}$: m/z 468.1407, found 468.1407.

(S)-benzyl 3-(6-methyl-1,3,5,7-tetraoxo-6,7-dihydropyrrolo[3,4-*f*]isoindol-2(1*H*,3*H*,5*H*)-yl)-2-(tritylamino)propanoate (3.19) Compound **3.15** (0.72 g, 1.6 mmol) was dissolved in DMF (23 ml) with diisopropylethylamine (1.3 ml, 7.4 mmol) added. 1,2,4,5-Benzenetetracarboxylic dianhydride (0.36 g, 1.6 mmol) was added followed by methylamine hydrochloride (0.22 g, 3.3 mmol). The reaction was stirred capped at room temperature for 2 hours. The mixture was then heated to 120° C overnight. After cooling to room temperature, the reaction was filtered and the filtrate concentrated under reduced pressure. Purification of the crude product by column chromatography (70:25:5 DCM/hexanes/ACN) gave compound **3.19** as a white solid (0.26 g, 25%). ^{1}H NMR ($CDCl_3$) δ 8.09 (s, 2H), 7.51 (m, 6H), 7.28-7.11 (m, 12 H) 7.02 (m, 2H), 4.55 (d, *J* = 12.4 Hz, 1H), 4.31 (d, *J* = 12.4 Hz, 1H), 4.17 (m, 1H), 4.14 (m, 1H), 3.77 (m, 1H), 3.24 (s, 3H), 2.94 (bs, 1H). ^{13}C NMR ($CDCl_3$) δ 172.23, 166.05, 165.62, 145.28, 136.99, 136.70, 134.61, 128.59, 128.23, 128.18, 127.95, 127.87, 126.54, 118.03, 71.16, 67.28, 54.95, 42.19, 24.39. HRMS (CI) calculated for $C_{40}H_{32}N_3O_6$ $(M+H)^{+1}$: m/z 650.2291, found 650.2292.

(S)-benzyl 2-(tert-butoxycarbonylamino)-3-(6-methyl-1,3,5,7-tetraoxo-6,7-dihydropyrrolo[3,4-*f*]isoindol-2(1*H*,3*H*,5*H*)-yl)propanoate (3.20) Compound **3.19** (0.26 g, 0.4 mmol) was dissolved in a mixture of DCM (6 ml), methanol (0.12 ml), and TFA (94 μl). The reaction was stirred for 10 minutes. The reaction was concentrated under reduced pressure. The residue was dissolved in DCM (4 ml) with diispropylethylamine (212 μl, 1.2 mmol) added. Di-tert-butyl dicarbonate (97 mg, 0.44 mmol) was added to the mixture and the reaction was stirred overnight under nitrogen gas. The reaction mixture was diluted with DCM (20 ml) and washed with 0.5 M citric acid (3 x 10 ml), saturated $NaHCO_3$ (3 x 10 ml), and water (10 ml). The organic layer was concentrated under reduced pressure. Purification of the crude product by column chromatography (70:25:5 DCM/hexanes/ACN) gave compound **3.20** as a white solid (0.142 g, 69%). ^{1}H NMR ($CDCl_3$) δ 8.22 (s, 2H), 7.33-7.30 (m, 5H), 5.35 (d, J = 8 Hz, 1H), 5.18 (s, 2H), 4.72 (q, J = 4.8 Hz, 1H), 4.20 (m, 1H), 4.16 (m, 1H), 3.25 (s, 3H), 1.32 (s, 9H). ^{13}C NMR ($CDCl_3$) δ 169.82, 166.39, 166.17, 137.48, 137.17, 134.93, 128.84, 128.81, 118.56, 80.60, 68.28, 52.53, 40.49, 28.33, 24.78. HRMS (CI) calculated for $C_{26}H_{26}N_3O_8$ $(M+H)^{+1}$: m/z 508.1720, found 508.1720.

(S)-benzyl 2-(tert-butoxycarbonylamino)-3-(6-methyl-1,3,5,7-tetraoxo-6,7-dihydropyrrolo[3,4-*f*]isoindol-2(1*H*,3*H*,5*H*)-yl)propanoic acid (3.21) Compound **3.20** (0.30 g, 0.59 mmol) was dissolved in DMF (6 ml). 10% Palladium on carbon (63 mg, 10 mol%) was added and the solution was purged with hydrogen gas. The reaction was stirred under a balloon of hydrogen for 5 hours at room temperature. The solution was filtered and the solid rinsed with DMF. The filtrate was concentrated under reduced pressure to afford compound **3.21** as a white solid (0.24 g, 99%). ^{1}H NMR (DMSO) δ

8.18 (s, 2H), 7.16 (d, J = 8.4 Hz, 1H), 4.28 (q, J = 6 Hz, 1H), 3.98-3.90 (m, 2H), 3.09 (s, 3H), 1.25 (s, 9H). ^{13}C NMR ($CDCl_3$) δ 171.15, 166.30, 165.99, 155.28, 137.24, 136.69, 78.36, 51.28, 27.93, 24.15. HRMS (CI) calculated for $C_{19}H_{20}N_3O_8$ $(M+H)^{+1}$: m/z 418.1250, found 418.1250.

N-α-Trityl aspartic acid dibenzyl ester (3.22) Aspartic acid di-benzyl ester p-toluenesulfonate salt (10 g, 20.6 mmol) was suspended in dry DCM (83 ml). Triethylamine (6.1 ml, 43 mmol) was added dropwise. The solution was cooled to 0° C and trityl chloride (6.0 g, 21 mmol) was added. The reaction was stirred at 0° C under nitrogen for 6 hours. The reaction mixture was diluted with additional DCM (50 ml) and extracted with 0.1 M citric acid (3 x 100 ml), saturated $NaCO_3$ (3 x 100 ml), and water (100 ml). The organic layer was dried over Na_2SO_4 and concentrated under reduced pressure. Purification of the crude product by column chromatography (9:1 hexanes/EtOAc) gave compound **3.22** as a white solid (9.9 g, 87%). ^{1}H NMR ($CDCl_3$) δ 7.71-7.69 (m, 5H), 7.47-7.31 (m, 15H), 5.23 (dd, J = 6.8, 12.4 Hz, 2H), 4.90 (dd, J = 22.8, 12 Hz, 2H), 4.01 (m, 1H), 3.27 (d, J = 9.2 Hz, 1H), 2.85 (dd, J = 10, 4.8 Hz, 1H), 2.70 (dd, J = 8, 7.2 Hz, 1H). ^{13}C NMR ($CDCl_3$) δ 172.88, 170.06, 145.47, 135.48, 135.14, 128.50, 128.24, 128.17, 128.03, 127.96, 127.93, 127.68, 126.29, 71.04, 66.60, 66.14, 53.47, 39.73. HRMS (CI) calculated for $C_{37}H_{34}NO_4$ $(M+H)^{+1}$: m/z 556.2488, found 556.2484.

N-α-Trityl serine lactone (3.23) Compound **3.22** (9.7 g, 17.5 mmol) was dissolved in dry toluene (175 ml) and cooled to -78° C under argon. A 1.5 M DIBALH in toluene solution (17.5 ml, 26 mmol) was added and the reaction was stirred for 2 hours at -78° C. The reaction was allowed to warm to 0° C over 3 hours. The reaction was stirred an

additional 18 hours at 4° C. The reaction was quenched with an aqueous solution of NH_4Cl (150 ml) and extracted with EtOAc (3 x 100 ml). The organic layer was removed under reduced pressure. Purification of the crude product by column chromatography (9:1 hexanes/EtOAc) gave compound **3.23** as a clear oil (3.9 g, 64%). ^{1}H NMR ($CDCl_3$) δ 7.64-7.62 (m, 5H), 7.35-7.23 (m, 10 H), 4.02 (t, *J* = 8.8 Hz, 1H), 3.68 (m, 1H), 3.44 (t, *J* = 8 Hz, 1H), 2.91 (bs, 1H), 1.57 (m, 1H), 1.18 (m, 1H). ^{13}C NMR ($CDCl_3$) δ 178.00, 145.40, 128.51, 127.85, 126.51, 70.91, 65.06, 53.00, 32.09. HRMS (CI) calculated for $C_{23}H_{22}NO_2$ $(M+H)^{+1}$: m/z 344.1651, found 344.1649.

N-α-*tert*-butyloxycarbonyl serine lactone (3.24) Compound **3.23** (1 g, 2.9 mmol) was dissolved in DCM (29 ml). Trifluoroacetic acid (0.67 ml, 8.7 mmol) was added and the reaction was stirred for 15 minutes at room temperature. The mixture was concentrated under reduced pressure. The residue was dissolved in DCM (29 ml) with diispropylethylamine (2.0 ml, 11.6 mmol) added. Di-tert-butyl dicarbonate (0.7 g, 3.2 mmol) was added to the mixture and the reaction was stirred overnight under nitrogen gas. The reaction mixture was washed with 0.5 M citric acid (3 x 10 ml), saturated $NaHCO_3$ (3 x 10 ml), and water (10 ml). The organic layer was dried over Na_2SO_4 and concentrated under reduced pressure. Purification of the crude product by column chromatography (7:3 hexanes/EtOAc) gave compound **3.24** as a white solid (0.34 g, 58%). ^{1}H NMR ($CDCl_3$) δ 5.40 (d, *J* = 6 Hz, 1H), 4.36 (t, *J* = 8.8 Hz, 2H), 4.16 (m, 1H), 2.59 (m, 1H), 2.14 (q, *J* = 10.8 Hz, 1H), 1.36 (s, 9H). ^{13}C NMR ($CDCl_3$) δ 175.42, 155.31, 80.07, 65.50, 49.79, 29.68, 28.03. HRMS (CI) calculated for $C_9H_{16}NO_4$ $(M+H)^{+1}$: m/z 202.1079, found 202.1082.

N-α-*tert*-butyloxycarbonyl homoserine benzyl ester (3.25) Compound **3.24** (0.65 g, 3.2 mmol) was dissolved in ethanol (11 ml). Sodium hydroxide (0.13 g, 3.2 mmol) was added to the solution and the reaction was stirred overnight at room temperature. The mixture was concentrated under reduced pressure. The residue was dissolved in dry DMF (4 ml). Benzyl bromide (0.39 ml, 3.2 mmol) was added to the solution and the reaction was stirred overnight at room temperature. The reaction mixture was diluted with EtOAc (25 ml) and washed with water (3 x 10 ml), $NaHCO_3$ (3 x 10 ml), and brine (10 ml). The organic layer was dried over Na_2SO_4 and concentrated under reduced pressure. Purification of the crude product by column chromatography (7:3 hexanes/EtOAc) gave compound **3.25** as a white solid (0.89 g, 89%). ^{1}H NMR ($CDCl_3$) δ 7.28 (m, 5H), 5.64 (d, *J* J= 8 Hz, 1H), 5.10 (d, *J* = 8 Hz, 2H), 4.44 (m, 1H), 3.62 (m, 2H), 3.39 (bs, 1H), 2.06 (m, 1H), 1.67 (m, 1H). ^{13}C NMR ($CDCl_3$) δ 172.54, 156.17, 135.11, 128.43, 128.27, 128.08, 80.15, 67.04, 58.13, 50.74, 35.52, 28.10. HRMS (CI) calculated for $C_{16}H_{24}NO_5$ $(M+H)^{+1}$: m/z 310.1654, found 310.1656.

(S)-benzyl 2-(*tert*-butoxycarbonylamino)-4-(5-methoxynaphthalen-1-yloxy)butanoate (3.26) According to the method used for **3.11**, compound **3.26** was obtained as a white solid in 72% yield. ^{1}H NMR ($CDCl_3$) δ 7.86 (d, *J* = 8.4 Hz, 1H), 7.81 (d, *J* = 8.8 Hz, 1H), 7.36 (m, 2H), 7.30 (m, 5H), 6.84 (d, *J* = 6.8 Hz, 1H), 6.75 (d, *J* = 7.6 Hz, 1H), 5.50 (d, *J* = 7.6 Hz, 1H), 5.15 (s, 2H), 4.66 (m, 1H), 4.18 (m, 2H), 3.98 (s, 3H), 2.50-2.37 (m, 2H), 1.45 (s, 9H). ^{13}C NMR ($CDCl_3$) δ 172.02, 155.28, 155.12, 153.87, 135.17, 128.45, 128.27, 128.20, 126.55, 126.41, 125.19, 124.96, 114.50, 114.13, 105.39, 104.47, 79.87, 67.15, 64.34, 55.43, 51.72, 31.72, 28.22. HRMS (CI) calculated for $C_{27}H_{31}NO_6$ $(M+H)^{+1}$: m/z 465.2151, found 465.2147.

(S)-benzyl 2-(*tert*-butoxycarbonylamino)-4-(5-methoxynaphthalen-1-yloxy)butanoic acid (3.27) Compound **3.26** (0.27 g, 0.58 mmol) was dissolved in DMF (6 ml). 10% Palladium on carbon (32 mg, 5 mol%) was added and the solution was purged with hydrogen gas. The reaction was stirred under a balloon of hydrogen for 5 hours. The palladium was filtered and washed with DMF and the filtrate was concentrated under reduced pressure. The product was dissolved in a small volume of DCM and precipitated with hexanes. Filtration of the precipitate gave compound **3.27** as a white solid (0.22 g, 99%). ^{1}H NMR (CD_3OD) δ 7.80 (d, *J* = 8.4 Hz, 1H), 7.75 (d, *J* = 8.4 Hz, 1H), 7.34-7.23 (m, 3H), 6.83 (t, *J* = 7.2 Hz, 2H), 4.47 (m, 1H), 4.17 (m, 2H), 3.90 (s, 3H), 2.43 (m, 1H), 2.21 (m, 1H), 1.38 (s, 9H). ^{13}C NMR (CD_3OD) δ 156.48, 155.53, 129.82, 129.26, 128.74, 128.02, 127.97, 127.92, 127.89, 126.26, 126.08, 115.32, 106.50, 105.53, 80.58, 65.74, 55.98, 52.55, 36.92, 28.71. HRMS (CI) calculated for $C_{20}H_{25}NO_6$ $(M+H)^{+1}$: m/z 375.1682, found 375.1685.

(S)-4-amino-2-(benzyloxycarbonylamino)butanoic acid (3.28) N-α-L-glutamine (5 g, 17.8 mmol) was dissolved in a EtOAc (23 ml), ACN (23 ml) and water (12 ml) mixture. (Diacetoxyiodo)benzene (6.9 g, 21.4 mmol) was added and the solution was stirred at 16° C for 30 minutes. The reaction was warmed to 20° C and stirred for 7 hours. The reaction mixture was concentrated under reduced pressure. The residue was dissolved in water (200 ml) and washed with diethyl ether (3 x 150 ml) and EtOAc (3 x 150 ml). The aqueous layer was concentrated and the residue was recrystallized from hot ethanol and diethyl ether to afford compound **3.28** as a white solid (2.3 g, 51%). ^{1}H NMR (D_2O) δ 7.23 (m 5H), 4.92 (d, *J* = 7.2 Hz, 2H), 3.85 (dd, *J* = 3.2, 5.6 Hz, 1H), 2.85 (t, *J* = 7.6 Hz, 2H), 1.98 (m, 1H), 1.82 (m, 1H). ^{13}C NMR (D_2O) δ 157.93, 136.47, 128.90, 128.50,

127.85, 67.15, 54.08, 36.89, 29.91. HRMS (CI) calculated for $C_{12}H_{17}N_2O_4$ $(M+H)^{+1}$: m/z 253.1188, found 253.1193.

(S)-2-(benzyloxycarbonylamino)-4-(7-methyl-1,3,6,8-tetraoxo-7,8-dihydrobenzo[*lmn*][3,8]phenanthrolin-2(1*H*,3*H*,6*H*)-yl)butanoic acid (3.29) 1,4,5,8-naphthalenetetracarboxylic acid anhydride (2.4 g, 9 mmol) was suspended in water (450 ml) with sonication. A 1 M aqueous KOH solution (42 ml) was added and heated to 80° C until all the solids have dissolved. The solution was acidified to pH 6.4 with H_3PO_4 (~5 ml). Compound **3.28** (2.3 g, 9 mmol) was added and the pH was readjusted to pH 6.4 if needed. The reaction was heated to 110° C for 16 hours then cooled to room temperature and filtered. To the filtrate, concentrated HCl (~15 ml) was added until a white precipitate formed. The solution was stirred at 80° C for 30 minutes and cooled to room temperature. The solution filtered and the solid was washed with water and dried. The solid (3.9 g, 7.6 mmol) was dissolved in DMF (126 ml) and a 2 M solution of methylamine in THF (7.6 ml, 15.1 mmol) was added. The reaction was stirred capped for 3 hours then heated at 110° C uncapped for 2 hours. The reaction was diluted with $CHCl_3$ (200 ml) and washed with 1 M $KHSO_4$ (4 x 100 ml) and brine (100 ml). The organic layer was dried over Na_2SO_4 and concentrated under reduced pressure. Purification of the crude product by column chromatography (gradient from 95:5 DCM/MeOH to 85:15 DCM/MeOH) gave compound **3.29** as a tan solid (2.0 g, 43%). ^{1}H NMR (DMSO) δ 8.16 (s, 4H), 7.31 (m, 5H), 4.95 (s, 2H), 4.03 (m, 3H), 3.26 (s, 3H), 1.98 (m, 2H).

(S)-2-amino-4-(7-methyl-1,3,6,8-tetraoxo-7,8-dihydrobenzo[*lmn*][3,8]phenanthrolin-2(1*H*,3*H*,6*H*)-yl)butanoic acid hydrobromide (3.30) Compound **3.29** (2.0 g, 3.9 mmol) was suspended in 36% HBr in acetic acid (50 ml). The reaction was stirred for 1 hour at

room temperature. The reaction was diluted with diethyl ether (200 ml) and filtered. The solid was washed with diethyl ether and dried to afford compound **3.30** as a tan solid (1.65 g, 91%). ^{1}H NMR (DMSO) δ 8.45 (dd, *J* = Hz, 4H), 4.21 (m, 2H), 4.06 (m, 1H), 3.31 (s, 3H), 2.23 (m, 2H).

CHAPTER 4

Anchored Periplasmic Expression of High Affinity Peptides

4.1 CHAPTER SUMMARY

Introduction

In water, the intermolecular association of independent oligomers of NDI and DAN exhibits relatively high affinity considering the large charge repulsion between the strands (Gabriel 2002). It is proposed that this affinity could be exploited for use in protein expression tags by utilizing a peptide strand of natural aromatic amino acids, and possibly complimentary charge, in combination with a strand of synthetic NDI oligomers. This type of peptide tag for use in proteomics is conceptually analogous to the avidin-biotin system. This chapter describes the design and synthesis of fluorescently labeled and biotinylated NDI oligomers as well as the design and expression of an Anchored Periplasmic Expression (APEx) random peptide library in *E. coli*.

Goals

Utilizing a hybrid aromatic donor-acceptor system with a natural peptide and an abiotic oligomer is a new avenue to exploring the scope of donor-acceptor interactions in natural systems. Experiments in this chapter are designed to target the question: Can a natural peptide, possibly containing aromatic amino acids and complementary charge, bind an oligomer of NDI units with high affinity? The short term goal of this research was to design and screen a random peptide library using the APEx display system to determine its suitability for peptide display. These initial studies were intended to serve

as a platform for optimization of the expression system for selecting high affinity peptides for an NDI oligomer.

Approach

Two random 12-mer peptide libraries were expressed as the NlpA and gene III fusions in the APEx display system, in addition to several epitope tag fusion constructs, to determine the suitability of the original APEx system for peptide libraries. Both NDI monomers and tetramers were fluorescently labeled with fluorescein for use in screening the peptide libraries using fluorescence activated cell sorting (FACS). Biotinylated NDI tetramer was synthesized for use with phage display and surface plasmon resonance.

Results

Overall, the work in this chapter describes efforts to demonstrate the ability of the APEx system to be used for peptide display. The gene III fusion library expressed in *E. coli* gave no significant enrichment after two rounds of sorting by FACS with fluorescein-labeled NDI tetramer. Peptide expression and accessibility were probed using peptide epitope tags in place of the random peptide library. Both streptavidin-binding peptide and the FLAG peptide tag gave poor fluorescence enhancement over background signal. Combinations of both epitopes in opposite orientation in relation to the inner membrane anchor were expressed and gave only slight improvement in signal.

4.2 BACKGROUND

In the rapidly growing field of proteomics, the use of recombinant proteins has necessitated the utilization of a variety of techniques for selective purification, detection, and immobilization of individual proteins from complex mixtures of cellular extracts.

Many of these techniques involve recombinant hybrids, where proteins are expressed as fusions to polypeptide, called affinity tags. There have been two major classes of affinity tags; the use of small peptide tags or large proteins. Both systems have advantages and disadvantages. Small peptide tags are less immunogenic than large tags, should not interfere with protein folding and activity, and may not need to be removed. Due to their short sequences, however, achieving accessibility of the tag at a convenient location of attachment can be non-trivial. Large peptide tags can increase the solubility of the fused protein; however, they can interfere with protein folding and may need to be removed for future applications of the fused protein.

The choice of an affinity tag is not straightforward, due to the variety of conditions utilized for detection and purification. For purification, each affinity tag is adsorbed to a matrix using a specific binding partner and specific buffer conditions; alternate conditions are then used to elute the adsorbed fusion. The protein of interest must be stable to the conditions of both adsorption and elution. Several examples of these tags and their conditions are given in Table 4.1. (Terpe 2003)

Affinity Tag	Sequence	Matrix	Elution Conditions
His6	HHHHHH	Ni(II)-NTA	Imidazole (20-250 mM)
c-myc	EQKLISEEDL	monoclonal antibody	low pH
Strep-tag II	WSHPQFEK	modified streptavidin	2.5 mM desthiobiotin
FLAG	DYKDDDDK	anti-FLAG antibody	pH 3.0 or 2-5 mM EDTA
Calmodulin binding peptide (CBP)	KRRWKKNFIAVSAAN RFKKISSSGAL	Calmodulin	EGTA
Glutathione S-transferase (GST)	Protein (211 residues)	Glutathione	2-10 mM reduced glutathione
Hemagglutinin (HA)	YPYDVPDYA	anti-HA antibody	HA peptide
Maltose binding protein (MBP)	Protein (396 residues)	cross-linked amylose	10 mM maltose

Table 4.1 Representative affinity tags and conditions used for protein purification.

In addition to affinity tags that utilize two polypeptide partners, there is a class of affinity tags that utilize a polypeptide with a non-peptidic partner. The most common affinity tag of this type is the avidin-biotin system. Avidin-biotin is the most common system for attaching proteins to surfaces in protein arrays. Biotin can easily be reacted with surface exposed lysine residues in proteins using solution or solid-phase methods. (Wilchek 1990) Avidin can be purchased as a conjugate to a variety of surfaces including agarose, magnetic beads, glass beads, and gold nanoparticles. Avidin is a tetrameric protein with each subunit capable of binding one molecule of biotin (Figure 4.1). Since avidin is highly glycosylated and prone to non-specific surface binding, the prokaryotic version, streptavidin, is often used instead. Streptavidin is expressed from *Streptomyces avidinii* and in combination with biotin has a dissociation constant of 10^{-15}

M. While it is beneficial for an affinity tag application that this dissociation constant is the strongest known for any small molecule-protein complex, the drawback is that the resulting conditions for dissociation are very harsh and often require denaturing or extremely basic conditions. Additionally, mammalian cells contain biotin-dependent enzymes and the endogenous biotin produces significant background in detection applications and requires the use of reagents that block endogenous biotin. (Invitrogen Corp. 2005) There are also drawbacks arising from the proteinaceous nature of avidin. Avidin cannot be expressed as a fusion and must be stored in cold, aqueous conditions to avoid denaturing and prolong shelf-life. The use of a non-biological capture agent offers the advantage of protease-resistance, chemical, and structural stability.

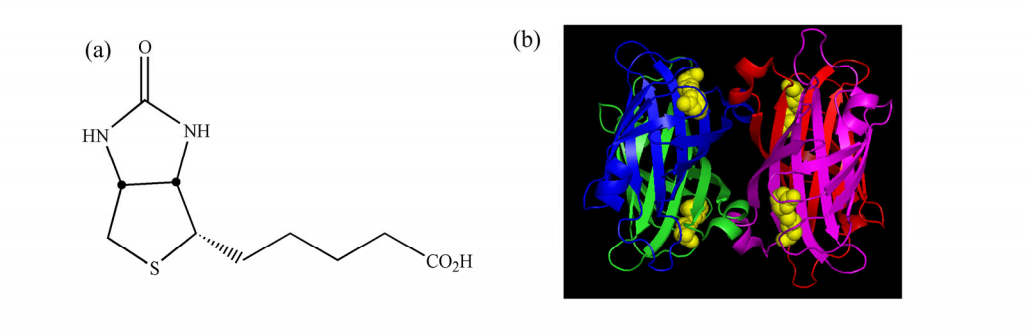


Figure 4.1 (a) Chemical structure of biotin. (b) Crystal structure of streptavidin with 4 biotins bound (PBD ID: 1STP).

Aromatic donor-acceptor interactions lend themselves to use in biological systems due to their high affinity in water. The original heteroduplex formation of NDI and DAN tetramers has an association constant of 3.5×10^5 M^{-1} that is not strong enough for use as an affinity tag. (Gabriel 2002) However, it should be possible to utilize longer oligomers with complementary charge to increase the association constant, affording a high affinity tag. Whereas the original heteroduplex contained two abiotic strands, it should be possible to replace the strand of DAN oligomers with a polypeptide that contains natural

aromatic electron-rich amino acids to obtain a peptide/artificial hybrid system of high affinity and selectivity. The peptide partner can then be expressed as a fusion, with a target protein for proteomics applications. The NDI oligomer can be chemically attached to a variety of surfaces and supports, which could be stored dry and at room temperature without denaturing or susceptibility to protease cleavage. (Figure 4.2)

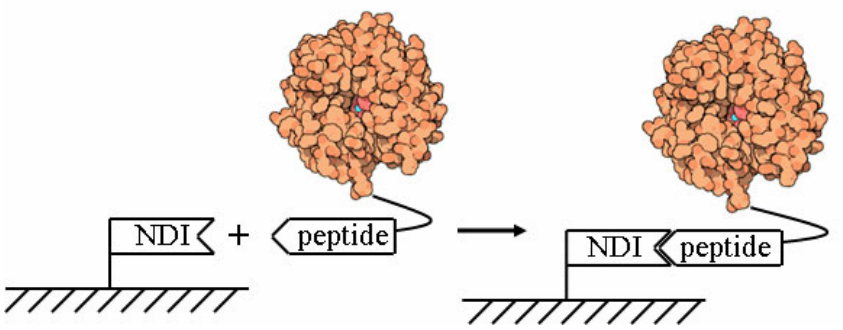


Figure 4.2 NDI-peptide affinity tag for surface attachment of proteins

Replacing the DAN oligomer with a designed peptide that binds an NDI oligomer with high affinity requires determination of an optimal peptide length and sequence. Due to the number of variables and choices in amino acid composition, rational design of the peptide was not feasible. Instead, the approach chosen involved screening a random peptide library. The work described in this chapter details the design and screening of a random peptide library and progress toward optimization of the APEx display system for peptide libraries.

4.3 Library Design and Probe Synthesis

4.3.1 Anchored Periplasmic Expression (APEx)

Several factors were considered in the choice of a peptide display system. Phage display systems are the "gold standard" for peptide display libraries and several random

peptide libraries are commercially available. However, phage display has several drawbacks, including the requirement of attachment of the target molecule to solid support, and dependence on factors such as surface characteristics, avidity effects, and hydrodynamic conditions. (Francisco 1993) Instead, the method of library display chosen was Anchored Periplasmic Expression (APEx). APEx is a patented and licensed technology developed by the Georgiou/Iverson lab for antibody and protein library screening. (Harvey 2004) APEx utilizes some of the machinery of filamentous phage display to anchor a peptide in the inner membrane of *E. coli*.

The peptides are most commonly displayed on the filamentous phage via fusion to the *N*-terminus of the gene III minor coat protein (g3p). During phage morphogenesis in *E. coli*, g3p becomes transiently attached to the inner membrane via its *C*-terminus before becoming incorporated on the growing phage. While attached to the inner membrane, the peptide is displayed in the periplasm. (Barbas 2001) Previous work in the Iverson/Georgiou lab has shown that by using a phagemid that encodes a C-terminal antibody-g3p fusion, an antibody library can be displayed in the periplasm of *E. coli*. Additionally, the use of the phagemid in *E. coli* combined with infection by helper phage allows the creation of phage libraries. (Harvey 2004)

In addition to the C-terminal antibody fusion, an alternate N-terminal fusion system was designed using the NlpA lipoprotein targeting motif. NlpA is secreted into the periplasm where a diacylglyceride group is attached to an N-terminal cysteine residue. The signal peptide is cleaved, the modified cysteine is fatty acylated, and the fatty acid inserts into the inner membrane. (Mingarro 1997) In the APEx system, the leader peptide and the first six amino acids of the mature NlpA are encoded as a fusion to an expressed protein of interest. The six amino acids of the NlpA protein contain the acylation site as well as the inner membrane targeting site. (McGregor 1994)

In the APEx method (shown in Figure 4.3), the fusion is expressed and attached to the inner membrane. The outer membrane is then disrupted via osmotic shock and the displayed peptide can bind an exogenous, fluorescently labeled probe. The fluorescently labeled spheroplasts are then sorted by flow cytometry and the DNA of isolated cells is rescued by PCR. After several rounds of sorting, the DNA can be sequenced to obtain the sequence of the displayed peptide. This method takes advantage of the relatively smooth surface of the inner membrane, as compared to display on the outer membrane of Gram-negative bacteria, where the presence of lipopolysaccharides can prevent the binding event due to steric hindrance. (Harvey 2004)

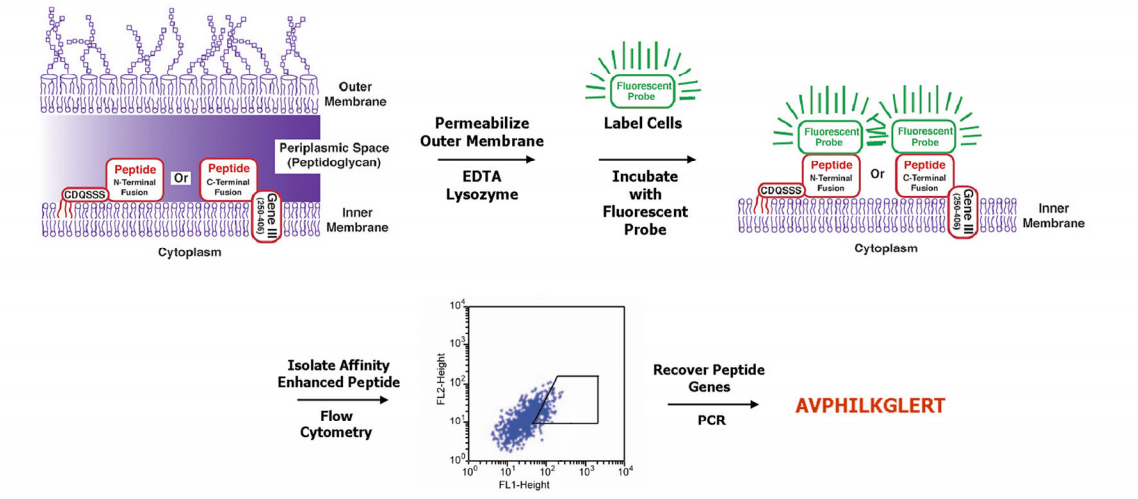


Figure 4.3 Anchored periplasmic expression system

As mentioned, the fluorescently labeled spheroplasts are sorted by flow cytometry. Fluorescent Activated Cell Sorting (FACS) uses an argon laser to measure physical characteristics of cells, such as size and shape, through detection of forward and side light scattering in addition to fluorescence. The cells can then be sorted either mechanically or electrostatically based on fulfillment of user-defined criteria for both size

and fluorescence intensity. With sorting rates up to 100,000 cells per second, screening a library of 10^6-10^7 cells becomes feasible. (Orfao 1996)

4.3.2 Library Construction

While it is not obvious what length of peptide is required for binding to an NDI oligomer, a 12-mer peptide library was chosen as the starting point in the library construction. The library was constructed with both the NlpA and g3p fusions via the same process. Initially, the library construction was attempted using a single long oligonucleotide composed of the entire forward strand between the restrictions sites of the vector and two small PCR primers.

PCR was used to make and amplify the double stranded DNA. While forming the double stranded DNA in the first round of PCR was successful, the amplification did not occur cleanly due to mis-annealing in the randomized region. This gave rise to different sized products, visible in agarose gels as fuzzy, broad bands. Alternately, the g3p fusion was constructed using the pAK200 vector created by Krebber as shown in Figure 4.4, and the library was inserted into the vector in a procedure analogous to a *lac* repressor peptide library developed by Affymax. (Krebber 1997, Schatz 1996) The vector was digested with SfiI to remove the tetracycline resistance genes. The random library was inserted by utilizing the non-compatible sticky-ends of the SfiI sites with 3 synthetic oligonucleotides. The long forward strand (VJD-N12) contained the complete sequence between the SfiI sites, including the last 11 nucleotides of the pelB leader sequence, the random peptide sequence $(NNN)_{12}$, and the 12 nucleotides encoding a 4 amino acid linker between the random peptide and the g3p sequence. The two short reverse strands (VJD14 and VJD10) correspond to the non-random regions of the reverse strand. All three oligonucleotides were 5' phosphorylated. The oligos were annealed to the vector and ligated, and polymerase is used to form the complement of the randomized region as

shown in Figure 4.5. The nicked phagemid was then transformed into the Jude 1 strain of *E. coli*. Jude 1 was chosen due to its successful use in antibody APEx libraries. (Harvey 2004)

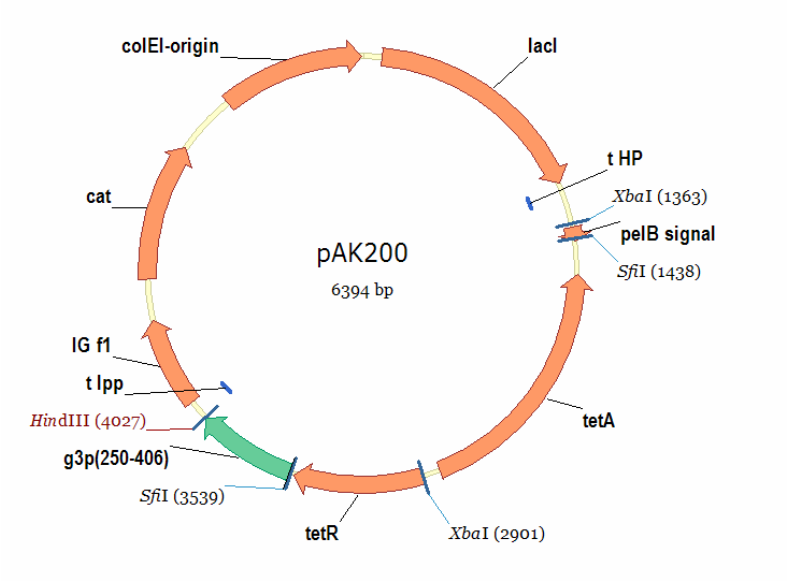


Figure 4.4 pAK200 vector map

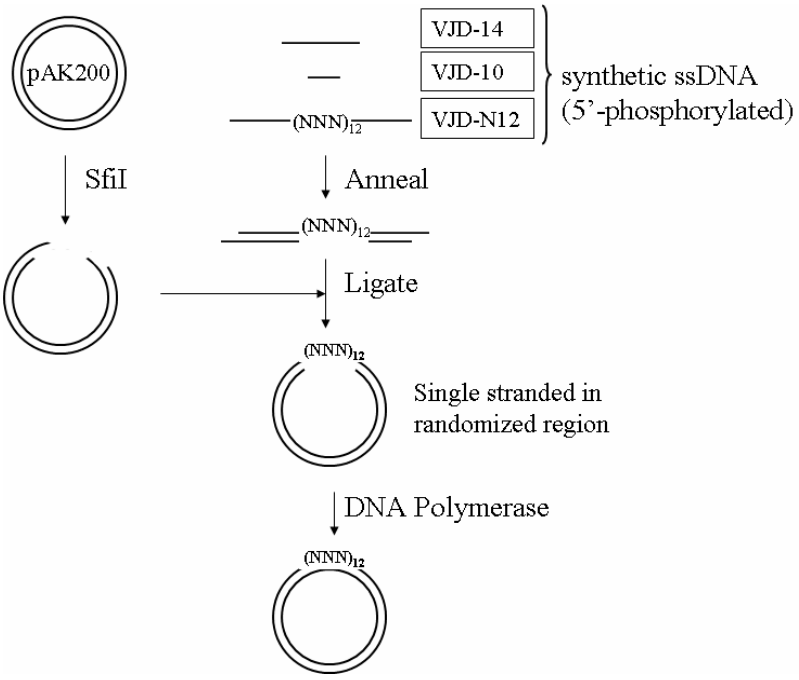


Figure 4.5 Vector construction

The library was also constructed as the NlpA fusion using the pAPEx vector shown in Figure 4.6 in the same manner as the g3p fusion. The tetracycline resistance genes as well as the His_6 and myc affinity tags were removed by double digestion of the vector with both SfiI and HindIII.

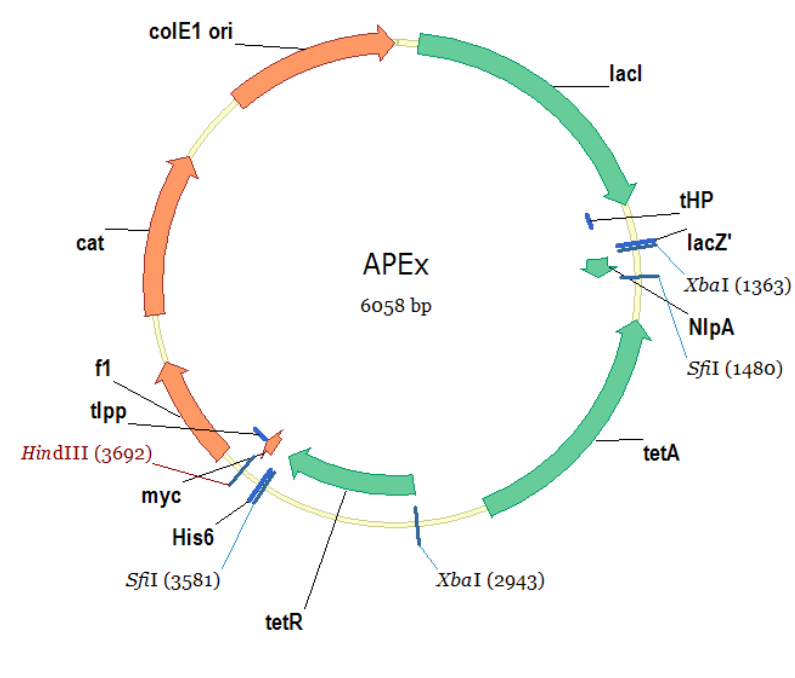


Figure 4.6 pAPEx vector map

4.3.3 Control Epitope Tags

In addition to the libraries, several controls were constructed to determine whether the peptides were being expressed and if a linker between the anchor and the library could improve binding accessibility. For both the g3p and NlpA fusions, a streptavidin binding sequence (SBS) (RLEICQNVCYYL) was cloned in place of the random library to determine if there was adequate expression. The SBS can be detected with a streptavidin-R-phycoerythrin fusion with a dissociation constant of approximately 10 nM.

(Bassette 2004) In order to determine if a linker could improve binding accessibility, the NlpA construct was used with the streptavidin binding sequence from above, a FLAG epitope (DYKDDDK), a FLAG-SBS conjugate, the reverse orientation conjugate, SBS-FLAG, and the longer FLAG-FLAG-SBS conjugate. The FLAG epitope was detected using an anti-FLAG antibody labeled with fluorescein isothiocynate (anti-FLAG-FITC) with an approximate dissociation constant of 6 nM (Wegner 2002). All the control and epitope constructs were constructed by annealing the complete forward and reverse oligonucleotides to the digested plasmid followed by ligation. (Figure 4.7)

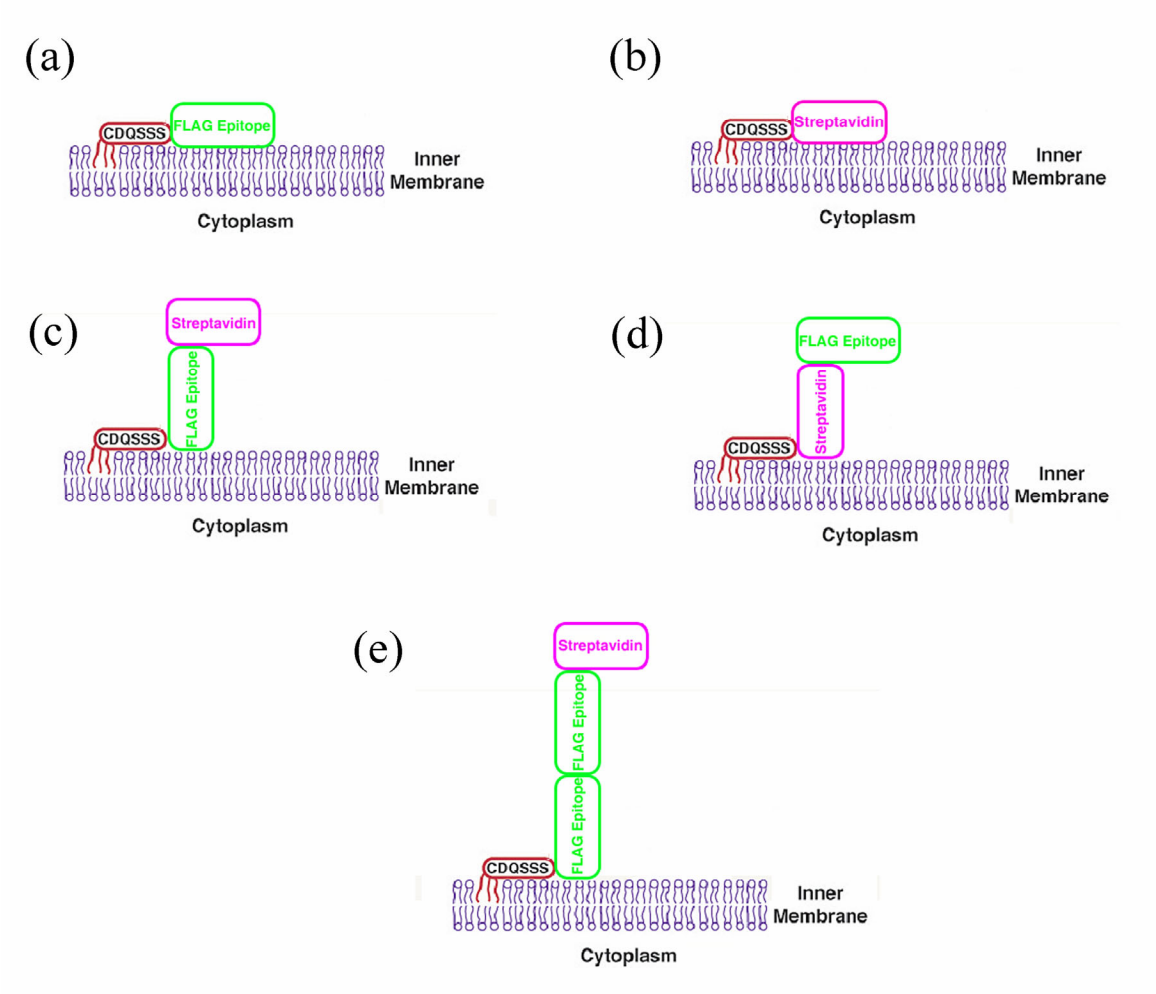


Figure 4.7 Control epitope tags for linker determination in the NlpA construct. (a) FLAG, (b) Streptavidin binding peptide (SBP), (c) FLAG-SBP, (d) SPB-FLAG, (e) FLAG-FLAG-SBP

4.3.4 Fluorescein and Biotinylated NDI Conjugates

The fluorescently labeled NDI tetramer was synthesized from Fmoc-Asp(OtBu)-NDI monomer using solid phase peptide synthesis as previously described for the heteroduplex formation. (Gabriel 2002) A potential problem in using the fluorescein conjugate of NDI is intermolecular and intramolecular quenching. While intermolecular quenching can be avoided by dilution, the problem of intramolecular quenching had to be

approached by using a small linker between the *N*-terminal NDI unit and 5,6-carboxyfluorescein. To determine the optimal linker length, three derivatives were synthesized with varying linker length: no linker **4.1**, β-alanine **4.2**, and di-glycine **4.3** (Figure 4.8). 5,6-Carboxyfluorescein was coupled to the N-terminus of the tetramer on 2-chlorotrityl chloride resin prior to resin cleavage using N,N-diisopropylcarbodiimide (DIC)/hydroxybenzotriazole (HOBt). (Fischer 2003) 2-Chlorotrityl chloride resin was chosen due to the mild conditions for cleavage.

An additional potential problem with fluorescein labeled NDI tetramer was self-stacking of the NDI units in the tetramer which would preclude peptide binding. For this reason, a fluorescently labeled NDI monomer with a di-glycine linker (**4.4**) was also synthesized via solid phase peptide synthesis.

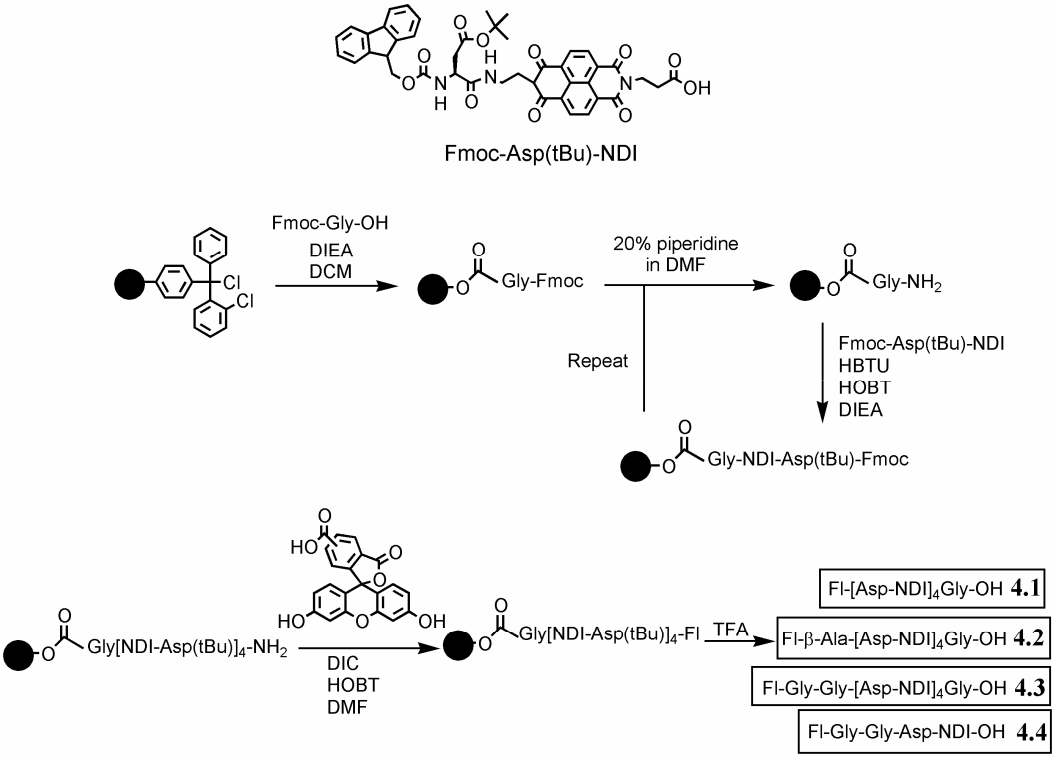


Figure 4.8 Solid phase peptide synthesis of fluorescently labeled NDI tetramers **4.1**-**4.3** and monomer **4.4**

With the need to measure association constants via surface plasmon resonance for sequences of interest, the biotinylated NDI tetramer **4.5** and monomer **4.6** (Figure 4.9) with the same di-glycine linkers were also synthesized by solid phase peptide synthesis in a similar manner to Figure 4.8. The NDI tetramer was biotinylated prior to resin cleavage by biotin succimidyl ester **4.7** in the presence of HOBt. Attempts to couple biotin directly were unsuccessful due to the insolubility of biotin in organic solvents. The biotin succimidyl ester was synthesized from biotin and disuccimidyl carbonyl (DSC). (Parameswaran 1990)

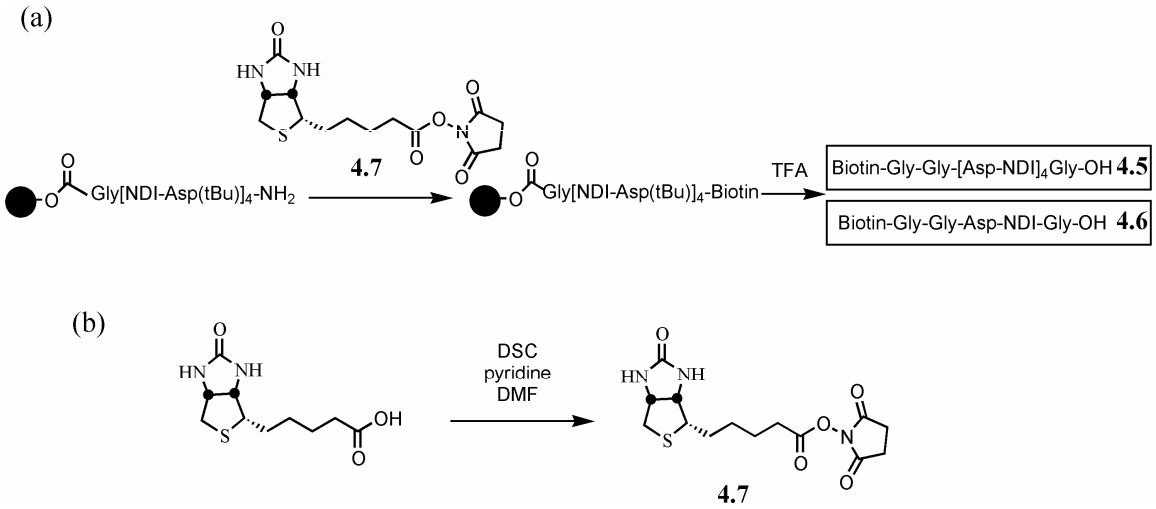


Figure 4.9 (a) Biotinylated NDI tetramer **4.5** and monomer **4.6**. (b) Synthesis of biotin succimidyl ester **4.7**.

4.4 RESULTS

4.4.1 Intramolecular Quenching

In order to determine the effect of intramolecular quenching, the concentrations of **4.1**, **4.2** and **4.3** were normalized via UV-Vis spectroscopy at λ_{max} for NDI at 384 nm and estimated to be approximately 40 μM using the extinction coefficient previously

determined for the NDI tetramer from Figure 1.x. (Gabriel 2002) The fluorescence emission spectrum for each of the three samples was compared as shown in Figure 4.10. While the difference in fluorescence between **4.1** and **4.2** at the maximum emission wavelength is within error, the fluorescence of the di-glycine linker **4.3** is approximately 2.5 times higher than either **4.1** or **4.2**. This corresponds to a quantum yield more than double that of no linker and a 4-atom linker. The fluorescent probe **4.3** containing the di-glycine linker was chosen for use in all further experiments.

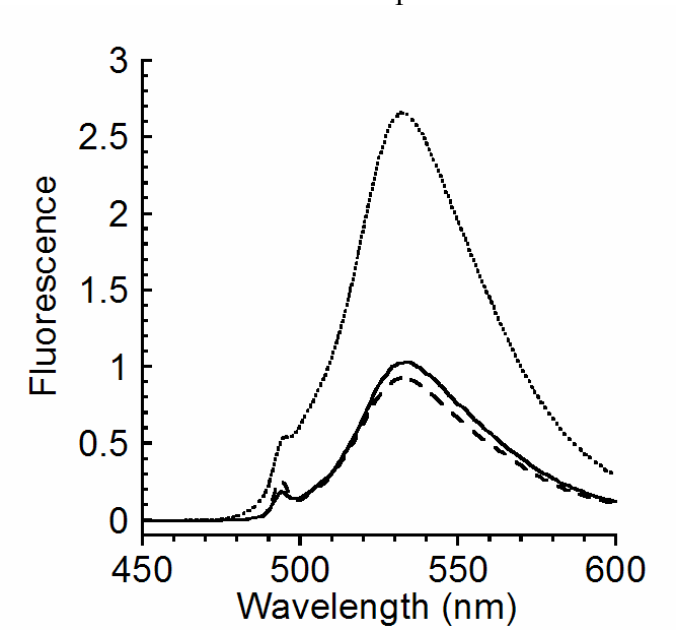


Figure 4.10 Fluorescence emissions of compounds **4.1** (—), **4.2** (- - -), and **4.3** (·····) at 40 μM in 100 mM sodium phosphate buffer, pH=7.0.

4.4.2 Non-Specific Binding

A potential problem that needed to be considered was non-specific binding of the fluorescent probe to the spheroplasts, which could result in false-positives during cell sorting. This was initially considered to be unlikely due to the electrostatic repulsion of the negatively charged surface of the inner membrane with the negatively charged probe. In order to test for non-specific binding, two strains of *E. coli*, MC4100 and Able C, that

did not contain phagemids for periplasmic expression, were grown and the outer membrane disrupted. The spheroplasts were incubated with fluorescent probe **4.3** at concentrations from 10 μM to 0.1 nM. The incubated spheroplasts were analyzed by FACS along with a negative control of spheroplasts that had not been incubated with the probe. The results are shown below in Figure 4.11. The non-specific binding threshold was determined to be >1 μM for MC4100 and ~1 μM for Able C. Since the concentrations required for the selection of high affinity clones are generally <10 nM, it was determined that non-specific binding was not a concern.

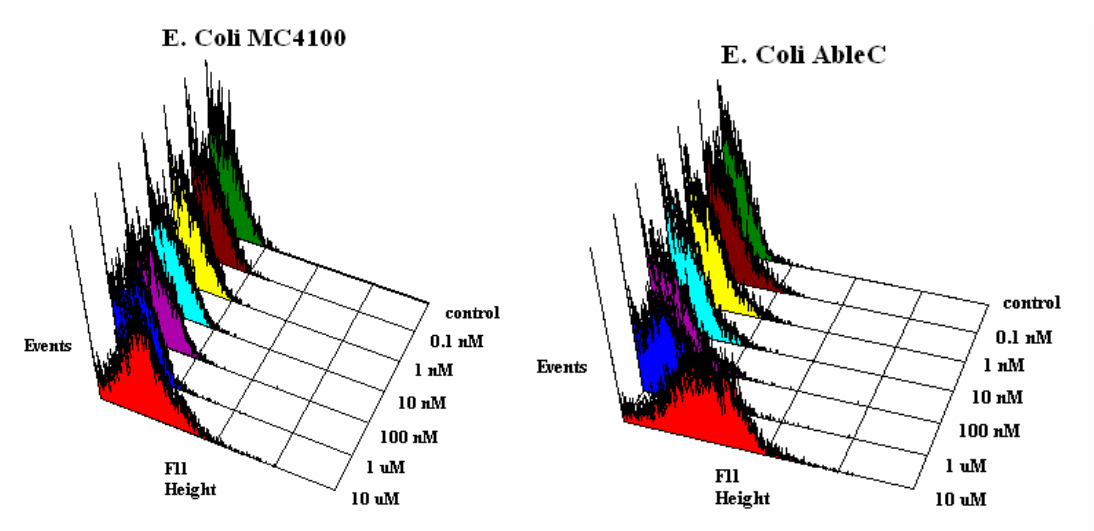


Figure 4.11 Non-specific binding of fluorescein labeled tetramer **4.3** at concentration from 0.1 nM to 10 μM with MC4100 and Able C strains of *E. coli*.

4.4.3 Random Peptide Library

The random library cloned into the g3p construct was transformed into the Jude 1 strain of *E. coli* to give approximately 10^7 clones as determined by dilution plates. The library cloned into the NlpA construct was also transformed into Jude 1 to give approximately $5x10^6$ clones. The g3p library was expressed and sorted at two

concentrations of fluorescein labeled tetramer initially: 1 μM and 1 nM. Figure 4.12 shows a sample of the population of the negative control and the library after one round of sorting. The sorted library at both concentrations does not show any significant increase in fluorescence over the negative control as seen by the arithmetic mean (M_n) of the population. The arithmetic mean for the 1μM sorted library was 14.67 compared to 14.68 for the negative control. The arithmetic mean for the 1 nM sorted library was 6.10 versus 6.22 for the negative control.

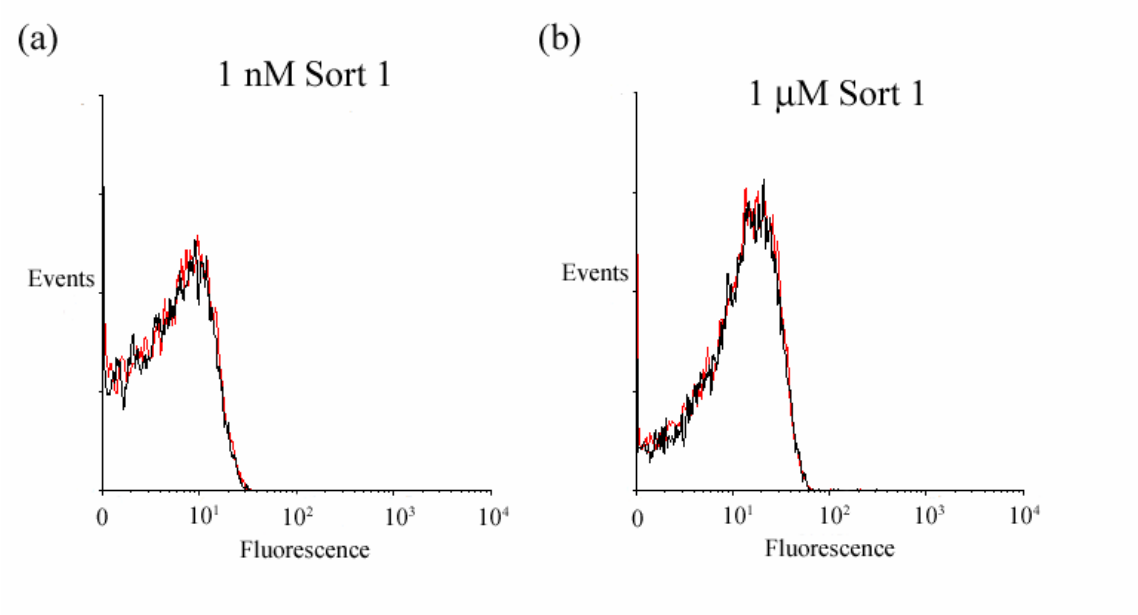


Figure 4.12 Histogram of first round sorting of the g3p 12-mer random peptide library at (a) 1nM and (b) 1 μM concentration (—) versus negative control (—).

After the first round of sorting, the randomized region of DNA from the isolated spheroplasts was amplified using rescue PCR, ligated into the pAK200 vector, transformed, and a second round of sorting was performed. For the clones isolated from the 1 μM and 1 nM first sort, the second round of sorting was completed after labeling again with 1 μM fluorescent tetramer. The second round of sorting of both populations gave a small two-fold shift to higher fluorescence over the negative control as shown in

Figure 4.13. The M_n of the 1 nM population at 1 μM concentration was 6.39 compared to 2.92 for the population prior to the second round of sorting. The M_n of the 1 μM population at 1 μM concentration was 5.52 versus 2.85 for the population prior to the second round of sorting.

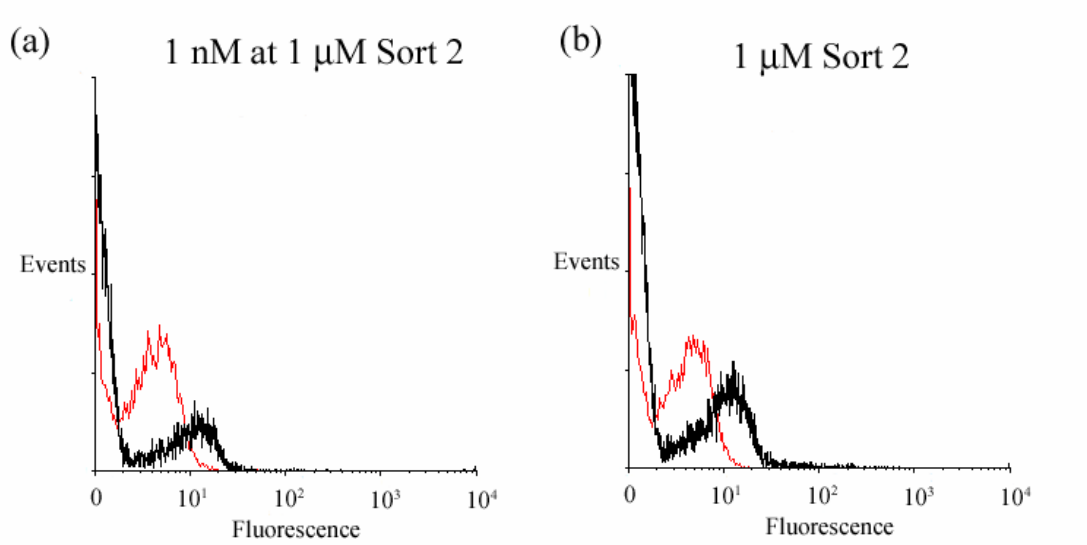


Figure 4.13 Histograms of a sample population from the second round of sortingof the g3p library: (a) clones from the 1 nM sort at 1 μM concentration of tetramer **4.3** (—)versus the population prior to the second round of sorting (—) (b) clones from the 1 μM sort at 1 μM concentration (—) versus the population prior to the second round of sorting (—).

Upon transformation of the rescue PCR product from the second round of sorting, 10 colonies were picked from each population for screening. The three highest fluorescent clones from the second round of sorting of the 1 nM population at 1 μM and the 1 μM population at 1 μM concentration are shown in Figure 4.14. From the 1 nM at 1 μM sort, the mean for the negative control is 2.99 versus 7.72, 6.13, and 6.20 for the three highest clones. From the 1 μM sort, the M_n for the negative control is 4.83 versus 6.96, 7.41, and 5.71 for the three most fluorescent clones. The results of the clones are consistent with sample populations from Figure 4.13.

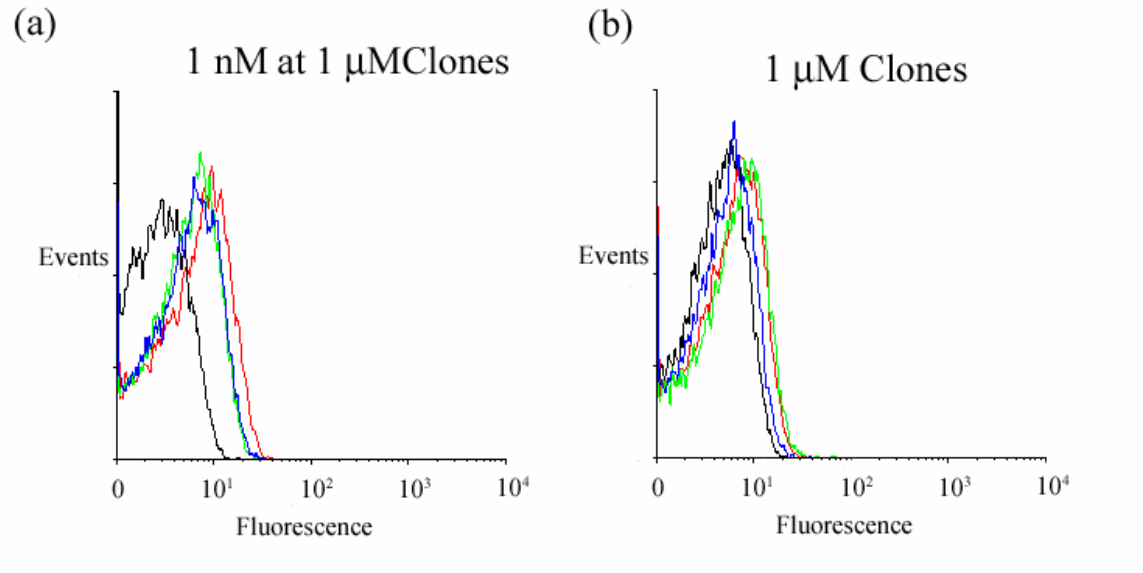


Figure 4.14 Histograms of the three highest fluorescing colonies from the g3p library for (a) the 1 nM at 1 µM second sort: (—) 15-9, (—) 15-3, (—) 15-4, and (—) negative control. (b) the 1 µM second sort: (—) 16-4, (—) 16-8, (—) 16-9, and (—) negative control.

While it is standard procedure to sort antibody libraries from PBS solutions, the large concentration of ions can interfere with electrostatic interactions due to screening of the charges by the high salt content of PBS. (Olsen 2000) As an alternate, isotonic 1% sucrose solution, known to avoid plasmolysis that is problematic with other low-ionic strength buffers, was compared to PBS in a single round of sorting of the g3p library labeled at 1 µM concentration of tetramer as shown in Figure 4.15. While the population sorted from PBS showed a shift to higher fluorescence (M_n= 84.99) over the unsorted library (M_n= 69.86), the population sorted from 1% sucrose showed a decrease in fluorescence from the unsorted library.

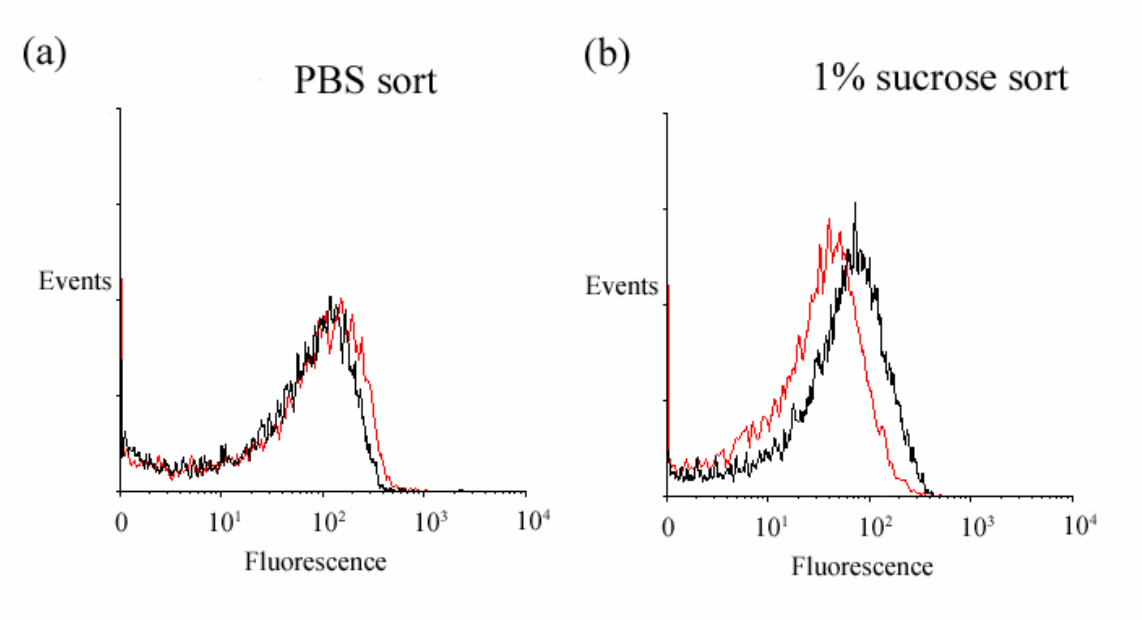


Figure 4.15 Histograms of the g3p random peptide library (—) compared to the isolated population sorted from 1 μM **4.3** (a) PBS and (—) (b) 1% sucrose (—).

4.4.4 Control Epitope Constructs

While it has been shown that both the g3p and NlpA constructs are effective for antibody libraries, a direct comparison of the two display constructs with peptides was needed. (Harvey 2004) The streptavidin binding peptide (SBP) was cloned in both pAK200 and pAPEx vectors. The spheroplasts were labeled with 50 nM streptavidin-R-phycoerythrin (SAPE) in PBS as well as blank Jude 1 spheroplasts as a negative control as shown in Figure 4.16. Both the NlpA and g3p constructs showed similar shift to higher fluorescence. With no clear benefit for either display system, NlpA was chosen to be used for the remainder of the epitope constructs.

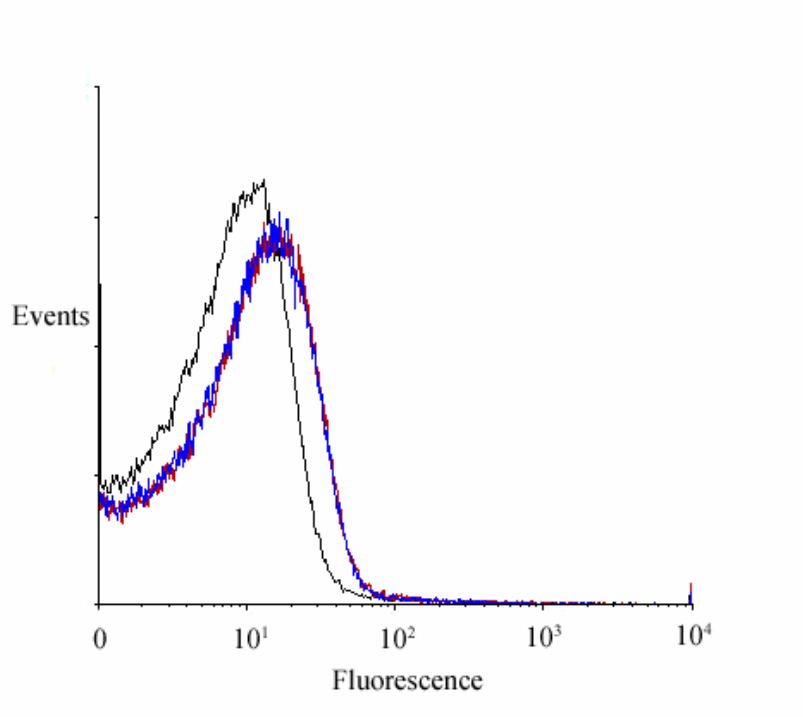


Figure 4.16 Histogram of streptavidin binding peptide (SBP) labeled with 50 nM streptavidin-R-phycoerythirin (SAPE) in both the g3p (—) and NlpA (—) constructs compared to a negative control (—).

In order to test whether a linker between the anchor peptide and the peptide of interest would improve binding accessibility, two epitope constructs were created using a single FLAG epitope peptide as the linker and two FLAG peptides as the linker for the SBP sequence. Figure 4.17 shows the histogram comparing no linker, a single FLAG linker, and a FLAG-FLAG linker compared to a negative control for binding of SAPE to the SBP sequence at 100 nM concentrations. While the FLAG-FLAG linker is the longest at 14 amino acids, the single FLAG linker has the highest mean fluorescence at double the value for the negative control. The mean fluorescence for no linker and the FLAG-FLAG linker are the same as the negative control.

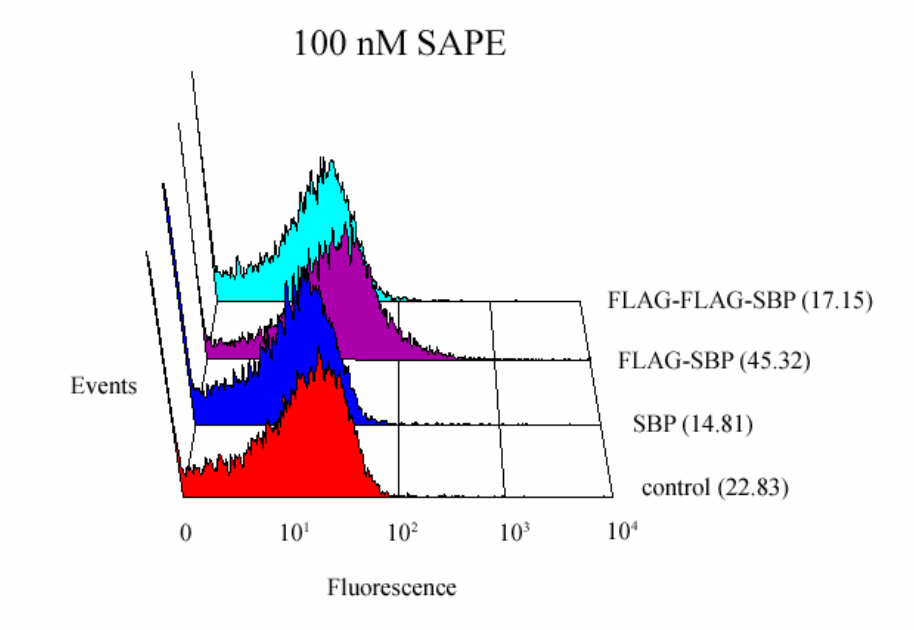


Figure 4.17 Histogram of NlpA-SBP fusion labeled with 100 nM SAPE with no linker, a single FLAG linker, and a FLAG-FLAG linker compared to a negative control. M_n shown in parentheses.

To determine if the FLAG epitope tag is also accessible to monitor expression when used as a linker, the FLAG-SBP and FLAG-FLAG-SBP conjugates were compared with the FLAG epitope displayed without the SBP sequence. The conjugates were labeled with 5 μg per mL of an FITC labeled anti-FLAG antibody in addition to a negative control and screened as shown in Figure 4.18. While all FLAG epitope constructs showed increased mean fluorescence over the negative control, the FLAG-SBP conjugate showed the highest mean fluorescence with a 5-fold increase. In addition, the increase in mean fluorescence for the FLAG epitope labeling was higher across all conjugates than for labeling the SBP epitope.

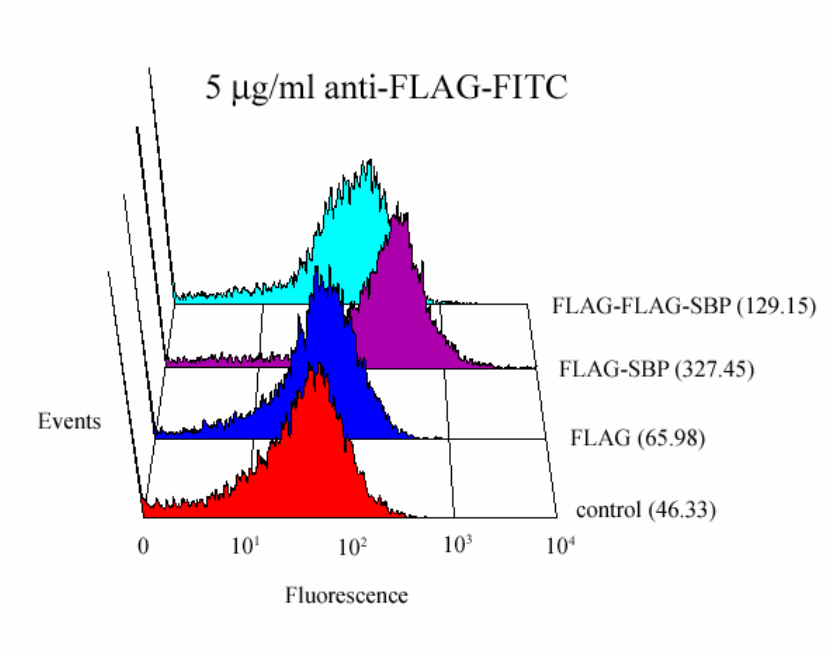


Figure 4.18 Histogram comparing labeling of Nlpa-FLAG fusion with the FLAG epitope used as a linker in the FLAG-SBP and FLAG-FLAG-SBP and a negative control. All fusions are labeled with 5 µg/mL FITC labeled anti-FLAG antibody. M_n shown in parentheses.

Due to the enhanced labeling of the FLAG-linked SBP sequence, a conjugate was made where the position of the FLAG and SBP epitopes were reversed relative to the anchor sequence (Figure 4.7). The FLAG-SBP and SBP-FLAG conjugates were labeled with SAPE and the FITC labeled anti-FLAG antibody separately for comparison by flow cytometry as shown in Figure 4.19. While binding of the SBP sequence in all the conjugates showed no significant differences, binding to the FLAG epitope varied greatly. The largest mean fluorescence was seen with the FLAG-SBP conjugate, where the FLAG epitope is used as a linker between the anchor and the SBP sequence. Interestingly, the reversed conjugate with the FLAG epitope further from the membrane anchor displayed a smaller increase in mean fluorescence; only 3.5 times that of the

negative control compared with ~6 times increase for the FLAG epitope directly next to the anchor.

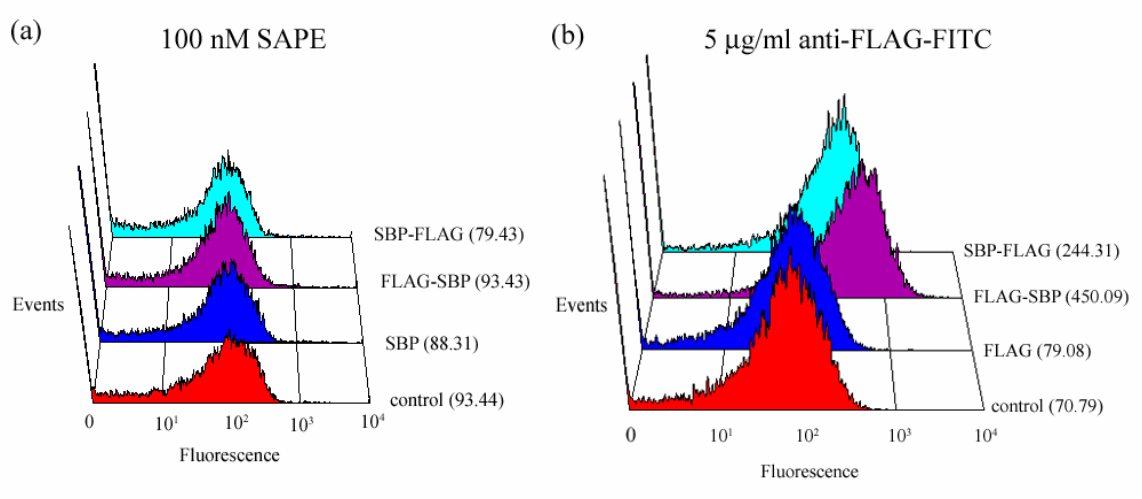


Figure 4.19 Histogram comparing labeling of the FLAG-SBP and SBP-FLAG conjugates with a negative control by (a) 100 nM SAPE and (b) 5 µg/mL FITC labeled anti-FLAG antibody. M_n shown in parentheses.

4.5 DISCUSSION

4.5.1 Random Peptide Library

Though both the NlpA and g3p 12-mer random peptide libraries were constructed, the sorting was performed with the g3p library due to the ability to transform the random peptide from the APEx system in *E. coli* directly to a phage display system. With no prior knowledge of how strongly a natural peptide could bind an oligomer, the sorting was performed using the fluorescently labeled NDI tetramer **4.3** at both 1 µM and 1 nM. From the non-specific binding experiment, it was determined 1 µM was the highest concentration that could be used for sorting without having to deconvolute non-specific

binding from the fluorescent signal. However, clones selected from a first round of sorting at 1 nM label concentration would provide only the highest affinity peptides.

After one round of sorting at both concentrations, the isolated populations did not show any enhancement in fluorescence or binding affinity. While disappointing, it was not completely unexpected since the highest 0.5% of fluorescent spheroplasts were collected. Considering there were 10^6 clones in the library, approximately 5,000 clones were collected. It is not unusual to perform at least 3 rounds of sorting to obtain only a few clones with the highest affinity. However, it was expected that after a second round of sorting, a significant shift in the mean fluorescence would be seen especially in the 1 nM clones from the first round that had been sorted in the second round at a label concentration of 1 μM. Results indicate only a two fold increase in mean fluorescence from the first round to the second round sorting. This was substantially less than the typical antibody libraries that show at least an order of magnitude or more increase in mean fluorescence over the original population. (Harvey 2004) In addition, the mean fluorescence was similar for both the 1 μM and 1 nM populations. One would expect the 1 nM clones to have a higher mean fluorescence if they had a higher binding affinity.

This small increase in fluorescence could be attributed to multiple variables. The first to be considered was weak binding due to screening of the charges by the high salt (100 mM) concentration of PBS. The rationale for the proposed interaction of an oligomer of NDI with a natural peptide was the strength of the aromatic donor-acceptor interaction as well as potential complementary charge. If the high salt concentration of PBS screens the charges on the NDI oligomer and peptide from "seeing" each other, the overall affinity would be significantly less. In a system of cell surface display of a protease of altered specificity with a fluorescent peptide probe, a high number of positive charges on the peptide localizes it to the highly negatively charged surface of the

membrane, thereby making the cell fluorescent for flow cytometry. (Olsen 2000) In labeling the system, 1% sucrose was used instead of PBS. Surprisingly, when the library was sorted from 1% sucrose, the resulting population had a lower mean fluorescence than the full library population. While the reduction seems significant, it is the likely result of an increase in the non-specific binding of the fluorescently labeled tetramer in 1% sucrose over PBS resulting in larger error margins.

Additionally, the ability of the tetramer of NDI to self-stack and thereby preculde binding to a peptide was considered. It was unclear whether a peptide could bind the NDI monomer with sufficient affinity to be sorted by flow cytometry due to a combination of dilution while resuspending the labeled spheroplasts and long sorting times with the potentially fast off-rate kinetics of weakly bound monomer.

4.5.2 Streptavidin Binding Peptide Control and Linkers

Instead of continuing to sort the g3p library with the NDI tetramer **4.3** or beginning to sort the library with the NDI monomer **4.4**, the direction of focus was changed in order to determine how well a known peptide in the APEx system could bind its target and optimize the APEx system for the known peptide. The knowledge could then be used to improve the random peptide library. The peptide chosen was the streptavidin-binding peptide (SBP) due to its 12 amino acid length. Also, the peptide was previously isolated from an alternate surface display library that utilizes the membrane protein OmpA to display peptides in an exposed outer loop region. (Bassette 2004) The peptide was expressed in both the *N*-terminal (NlpA) and *C*-terminal (g3p) fusions to determine if one display mechanism was better than another. As with antibodies, flow cytometry did not indicate a specific benefit to either system, although the increase in fluorescence over the negative control continued to be small in comparison to difference seen in antibody studies.

An additional potential cause for the small increase in fluorescence in the sorted library and the SBP peptide was peptide accessibility and preorganization. In phage display, the peptide is displayed at the end of a 406 amino acid protein. It has been shown that the full protein preorganizes the peptide in an extended conformation, usually significantly helical. In addition, there are three or more amino acids which act as a linker between gene 3 protein and the displayed peptide. This linker is extremely important in peptide accessibility. (Burritt 1996) While not a concern in an antibody library, the peptide in the NlpA fusion is only 6 amino acids from the fatty acid anchor. The FLAG epitope was chosen to be used as a possible linker for several reasons. First, the 7 amino acid sequence would essentially double the distance between the membrane anchor and the displayed peptide. Second, the FLAG epitope has an overall charge of -2 with a region of three consecutive negatively charged aspartate residues that should provide charge repulsion between the peptide and the membrane surface providing greater accessibility. Finally, the FLAG epitope could possibly be used to monitor peptide expression levels (i.e. an internal positive control).

To determine the effectiveness of the FLAG epitope as a linker for the SBP sequence, three constructs were compared: no FLAG linker, a single FLAG sequence linker, and two consecutive FLAG sequences. Interestingly, of the three constructs, the streptavidin (SAPE) had the largest mean fluorescence and highest affinity for the SBP sequence using the single FLAG sequence construct even though it was not the longest linker. Additionally, SAPE showed no binding to the linker-less and FLAG-FLAG linker constructs. It is unclear why there is no binding to either construct. While the single FLAG linker displayed a 2-fold increase in affinity and fluorescence, the expected value is at least an order of magnitude higher considering the binding constant is determined to be 10 nM and labeling was performed at 100 nM.

While SBP was still not accessible to binding by streptavidin, the FLAG epitope could still be used to monitor expression levels in this peptide display system. Comparison of the FLAG linker constructs as well as FLAG displayed alone showed and increase in mean fluorescence in all three constructs, albeit small when FLAG is displayed alone. The small increase in fluorescence when the FLAG epitope is displayed alone indicates the negative charges and the 6 amino acids in the NlpA sequence are not enough to provide accessibility to the displayed system. Interestingly, having an additional sequence displayed "outside of" (further from the membrane than) the FLAG sequence made it more accessible to the anti-FLAG antibody for binding where as having two FLAG sequences in a row reduced the accessibility.

Although it is not clear why the FLAG-SBP construct is the most accessible for both the streptavidin and anti-FLAG antibody, the results beg the question: Could displaying the FLAG linker "outside" the SBP sequence improve binding? The SBP-FLAG construct was designed to answer this question by reversing the order of the peptides relative to the position of the membrane anchor. While binding of streptavidin to the SBP sequences was not enhanced, the results did show the anti-FLAG antibody still has better access to the SBP-FLAG peptide than when FLAG is displayed alone. However, the fluorescence and accessibility is two fold higher when the SBP sequence is displayed "outside" the FLAG sequence. This suggests that FLAG epitope is an excellent tag for the APEx system for monitoring expression levels when displaying peptides.

One parameter to be considered in the analysis of the difference in binding of the anti-FLAG antibody and the streptavidin-R-phycoerythrin conjugate to their respective sequences is their size. Although the length of the FLAG and SBP are similarly small, there is a 3-fold difference in the sizes of the proteins used to detect them. The anti-FLAG antibody is approximately 80 kDa whereas the SAPE conjugate is a total of 292

kDa. Even though the binding affinity to their respective sequences is similar (6 nM versus 10 nM), it would make sense that the SAPE conjugate would have more accessibility problems strictly on the basis of size. It is likely that the SBP epitope requires a much longer and preorganizing linker than the FLAG sequence. However, it is clear from the experiments involving the FLAG and SBP epitopes expressed alone as conjugates to the NlpA sequence that a linker is absolutely required for binding. This is likely true for the random peptide library as well, even though the binding partner is significantly smaller at 1.7 kDa.

4.6 CONCLUSIONS

While the use of the APEx display system had been demonstrated for antibodies, the usefulness of the system with peptide libraries had been proposed but not demonstrated. Using three single-stranded oligonucleotides, a random 12 amino acid peptide library has been created as both the *N*-terminal NlpA fusion as well as the *C*-terminal g3p fusion. This general method can be used to create libraries of various sizes in *E. coli*. However, based on these results, the APEx system used for antibody libraries cannot be used as-is in the construction of peptide libraries. Peptides fused directly to the NlpA and g3p sequences are not readily accessible for binding. While the FLAG epitope sequence is accessible for monitoring expression levels, it is not necessarily the best choice as a linker for peptides displayed in this manner. The optimal choice of linker has not been determined, but a design will involve weighing complex factors such as length and preorganization ability. The APEx system with the FLAG epitope as a linker for a peptide library has the potential to be used as a screen for peptide epitopes with antibodies or for protein-binding sequences, however caution should be used with proteins or antibodies larger in size than approximately 100 kDa.

While APEx can be modified for use with a random peptide library, future work in finding a peptide that will bind a peptide with high affinity would benefit from the use of the traditional phage display methods utilizing the biotinylated NDI monomer and tetramer and infection of the g3p library with helper phage. Additionally, use of the β-hairpin system in Chapter 3 can be used to determine which amino acids have the highest affinity for the NDI monomer in order to make a targeted peptide library for screening.

4.7 EXPERIMENTAL METHODS

4.7.1 General Methods

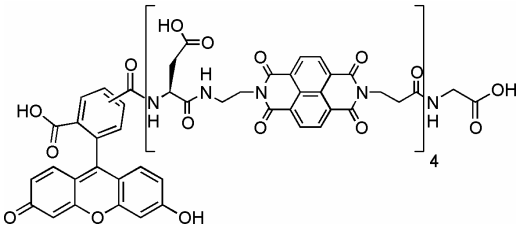
All oligonucleotides were purchased from Integrated DNA Technology (Coralville, IA) with desalting unless otherwise noted. Starting materials for chemical synthesis were purchased from Aldrich and used without further purification. Resin, coupling reagents, and amino acids were purchased from Novabiochem. Library sorting was performed on a DakoCytomation MoFlo flow cytometer (Carpinteria, CA) with excitation at 488 nm and detection for fluorescence and forward and side scattering at 530 nm. Scanning was performed on a Becton Dickinson FACScalibur (San Jose, CA) flow cytometer with excitation at 488 nm and detection at 530 nm (FL1) for fluorescein and 570 nm (FL2) for R-phycoerythrin. The collected data were processed in WinMDI 2.8 (http://facs.scripps.edu). Anti-FLAG-FITC antibody was purchased from Sigma and streptavidin-R-phycoerythrin was purchased from Invitrogen. All enzymes were purchased from New England Biolabs (Beverly, MA) and used with the supplied buffers except Sequenase 2.0 which was purchased from USB (Clevland, OH). UV-Vis spectra were taken of samples in 1 cm cuvettes on a Perkin-Elmer Lambda 35 spectrophotometer.

All sequencing of vectors was performed using the BRH06 primer (BRH06= GCGGATAACAATTTCAGACAG).

4.7.2 Synthesis of fluorescently labeled NDI compounds

Fmoc solid phase peptide synthesis

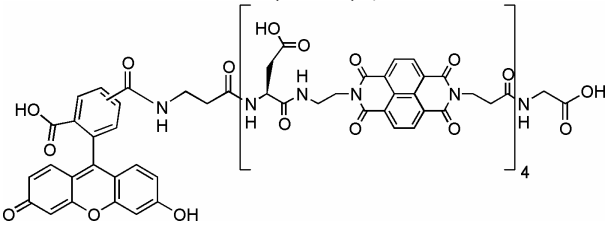
All solid phase peptide syntheses were performed on 2-chloro trityl chloride resin (100-200 mesh) with a loading of 1.6 mmol/g on a 0.25 mmol scale. Fmoc-Gly-OH (3 eq) was dissolved in dicholomethane (DCM) with diisopropylethylamine (DIEA, 3.3 eq) added. The solution was shaken with the resin for 5 minutes. An additional portion (6.6 eq) DIEA was added and the suspension was shaken for 2 hours. Methanol (0.33 mL) was added to the resin and shaken for 10 minutes. The solution was drained and the resin was washed 3 times with dimethylformamide (DMF). Deprotection of the Fmoc protecting group was afforded by shaking the resin with 20% piperidine in DMF for 20 minutes. The resin was then washed three times with DMF and DCM. Coupling of the Fmoc-Asp(tBu)-NDI-OH monomer was accomplished by dissolving the monomer (3 eq), O-(Benzotriazol-1-yl)-N,N,N′,N′-tetramethyluronium hexafluorophosphate (HBTU) (3 eq), HOBt (3 eq) and DIEA (6 eq) in N-methylpyrrolidone (NMP). The solution was shaken with the resin for 2 hours followed by washing of the resin three times with DMF and DCM. Deprotection/coupling cycles were performed to synthesize the remaining sequence on the resin. Cleavage from the resin was afforded using a cocktail of TFA/TIPS/H_2O (95:2.5:2.5) for 2 hours. The cocktail was evaporated via a stream of nitrogen gas and the product was precipitated with cold ether and filtered.



Fluorescein-[Asp-NDI]$_4$-Gly-OH 4.1

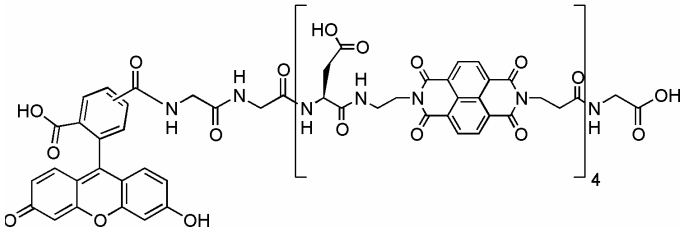
Fluorescein was conjugated to the *N*-terminus of the sequence on solid phase. (Fischer 2003) 5,6-Carboxyfluorescein (2.5 eq), diisopropylcarbodiimide (DIC, 2.5 eq) and HOBt (2.5 eq) were dissolved in DMF and added to the resin. The resin was shaken overnight and washed three times with DMF. To the resin, 20% piperidine in DMF was added and shaken for 45 minutes. After washing with DMF, an addition portion of 20% piperidine was added and shaken for 15 minutes. The resin was washed three times with DMF and DCM prior to cleavage.

After cleavage, the compound was purified via preparative HPLC with a gradient of 12 to 22 % B over 40 minutes, where solvent A was 10 mM ammonium acetate and solvent B was acetonitrile. Absorbance was monitored at 380 nM and the product eluted at 24 minutes. ESI MS: calculated 2346.5 $(M+H)^+$, found 2346.8.



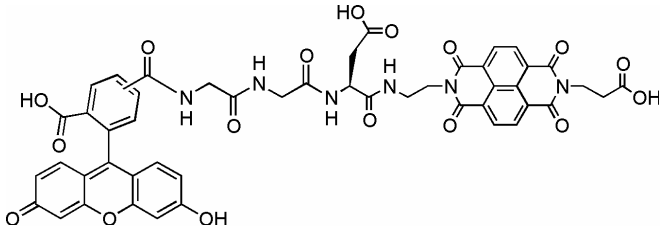
Fluorescein-β-Ala-[Asp-NDI]$_4$-Gly-OH 4.2

After cleavage, the compound was purified via preparative HPLC with a gradient of 12 to 22 % B over 40 minutes. The product eluted at 26 minutes. ESI MS: calculated 2418 $(M+H)^+$, found 2419.



Fluorescein-Gly-Gly-[Asp-NDI]$_4$-Gly-OH 4.3

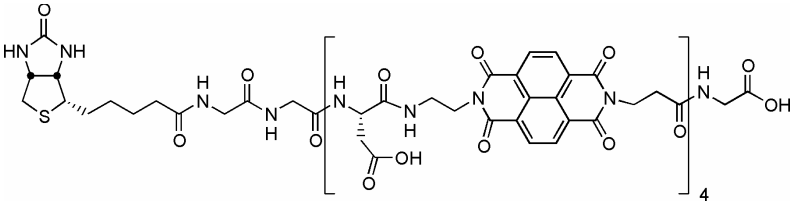
After cleavage, the compound was purified via preparative HPLC with a gradient of 12 to 22 % B over 40 minutes. The product eluted at 26 minutes. ESI MS: calculated 2462.1 $(M+H)^+$, found 2462.0.



Fluorescein-Gly-Gly-Asp-NDI-OH 4.4

After cleavage, the compound was purified via preparative HPLC with a gradient of 10 to 20 % B over 40 minutes. The product eluted at 13 minutes. MALDI MS: calculated 969.8 $(M+H)^+$, found 970.1.

4.7.3 Synthesis of biotinylated NDI compounds

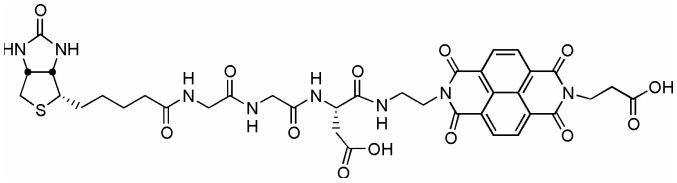


Biotin-Gly-Gly-[Asp-NDI]$_4$-Gly-OH 4.5

The same solid phase synthesis methods were used for the biotinylated compounds as described above. Conjugation of the biotin was achieved by dissolving biotin succimidyl ester **4.7** (5 eq), HOBt (5 eq), and DIEA (10 eq) in dry DMF and

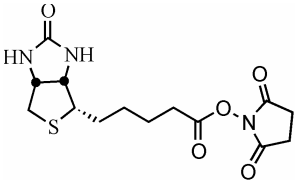
adding to the resin. The resin was shaken overnight and rinsed 3 times with DMF and DCM.

After cleavage, the compound was purified via preparative HPLC with a gradient of 12.5 to 32 % B over 45 minutes. The product eluted at 26 minutes. ESI MS: calculated 2350.6 $(M+Na)^+$, found 2349.7.



Biotin-Gly-Gly-Asp-NDI-OH 4.6

After cleavage, the compound was purified via preparative HPLC with a gradient of 12 to 25 % B over 39 minutes. The product eluted at 11 minutes. MALDI MS: calculated 837.8 $(M+H)^+$, found 837.4.



Biotin succimidyl ester 4.7

Biotin (2 g, 8.2 mmol) was suspended in dry DMF (25 mL). Dry pyridine (0.82 mL, 9.6 mmol) was added and the mixture was dissolved by stirring at 60° C. Disuccimidyl carbonate (2.185 g, 8.6 mmol) was added to the solution and was stirred at 60° C for 3 hours, then room temperature for 2 hours. The solvent was removed by rotary evaporation. To the solid, 35 mL of isopropanol was added and the suspension was cooled on ice. The solid was filtered and washed with isopropanol and ether. The solid was dissolved in 20 mL warm DMF and 80 mL of ether was added to precipitate the product. The product was isolated by filtration and dried in vacuo to yield 1.3 g (47%) of a white solid. ^{1}H NMR 300 MHz (DMSO) δ 6.45 (s, 1H), 6.38 (s, 1H), 4.32 (t, J = 6,

1H), 4.16 (m, 1H), 3.13 (m, 1H), 2.82 (s, 4H), 2.68 (t, J = 6.9, 2H), 2.61 (s, 1 H), 2.57 (s, 1H), 1.66 (m, 2H), 1.50 (m, 4H).

4.7.4 Intramolecular Quenching

Compounds **4.1**, **4.2**, and **4.3** were each dissolved in 100 mM phosphate buffer at pH=7 to a final concentration of 40 μM as determined by the absorbance at 364 nm. The excitation wavelength was 494 nm and the emission monitored between 450 and 600 nm with maximum emission at 530 nm.

4.7.5 Library Construction

G3p Library Vector Construction

The pAK200 vector was digested with SfiI for 2 hours at 50° C and purified by a 1% agarose gel. The three oligonucleotides used for construction of the random peptide library were VJD-N12 (CTGCTATGGCA$(NNN)_{12}$GGAGCAGCCTCGG), VJD14 (TGCCATAGCAGGCT), and VJD10 (AGGCTGCTCC). All three oligonucleotides were purchased as 5' phosphorylated and PAGE purified and were diluted to 5 nM stock solutions. To solution of 1 μL 5 M NaCl 2.5 μL 1 M Tris pH=7.4 and 43.5 μL water, 1 μL (50 pmol) each of VJD-N12, VJD14 and VJD10 were combined. The solution was heated to 70° C for 5 minutes then cooled slowly to 35° C. The mixture was then cooled to 4° C. A 2.6 pmol aliquot of the annealing mix was added to 1.3 pmol of the digested pAK200 vector, mixed on ice, and T4 ligase was added and the mixture was cooled to 17° C for 24 hours. To deactivate the T4 ligase, the mixture was heated to 65° C for 10 minutes followed by cooling to room temperature. To the mixture, 20 μL of a 2.5 mM solution of a mixture of deoxynucleotide triphophates was added followed by 1 μL of Sequenase 2.0 polymerase. The solution was incubated at 37° C for 1 hour. The vector was then purified via phenol-chloroform extraction and desalted by ethanol precipitation.

NlpA Vector Library Construction

The pAPEx vector was digested sequentially without intermediate purification with HindIII for 2 hours at 37° C followed by SfiI for 2 hours at 50° C. The three oligonucleotides used for construction of the random peptide library were VJB-N12-2 (CTGCTTCTAGT$(NNN)_{12}$TAGTGATACCGCA), VJBB17 (AGCTTGCGGTATC ACTA), and VJBF14 (ACTAGAAGCAGGCT). All three oligonucleotides were purchased as 5' phosphorylated and PAGE purified and were diluted to 5 nM stock solutions. The procedure for annealing, ligation, and polymerization were the same as described above for the g3p library.

Library Transformation

The vectors were transformed via electroporation in Jude 1 cells (*mcr*A, Δ(*mrr-hsd*RMS-*mcr*BC) ϕ80 d*lac*Z ΔM15 Δ*lac*X74 *deo*R *rec*A1 *end*A1 *ara*D139 Δ(*ara, leu*)7697 *gal*U *gal*K λ^- *rps*L *nup*G/F' $proAB^+$ $lacI^qZ\Delta M15$ Tn*10* Tet^r). Chloramphenicol was used for selection. Library size was determined from dilution plates.

4.7.6 Library Sorting

Induction and Spheroplast formation

A 20 mL culture in TB containing 2% glucose supplemented with chloramphenicol (30 μg/mL) was inoculated to an OD of 0.1 (~10 μL cell stock). The culture was shaked at 37° C for 2 hours and induced with 1 mM IPTG for 4 hours at 25° C. A 1 mL aliquot of the culture was removed and spun at 3000 rpm for 5 minutes. The broth was decanted and the pellet was resuspended in 350 μL 0.75 M sucrose-100 mM Tris pH=8. To the suspension, 35 μL of a 20 mg/mL solution of lysozyme was added. While slowly vortexing the suspension, 700 μL of 1 mM EDTA pH=7.5 was added dropwise. The suspension was gently shaken for 10 minutes at room temperature. Subsequently, 50 μL of 0.5 M MgCl2 was added and the mixture was gently shaken

again for 10 minutes on ice. The spheroplasts were pelleted by centrifugation for 10 minutes on maximum speed at 4° C. The supernatant was removed and the spheroplasts were resuspended in 500 μL of either 1xPBS or 1% sucrose.

Flow Cytometry Sorting

A 100 μL portion of spheroplasts was labeled in a total volume of 1 mL labeling solution and shaken at room temperature for 1 hour. The spheroplasts were pelleted at 13,200 rpm for 1 minute. The supernantant was decanted and the pellet was resuspended in either 1xPBS or 1% sucrose. The suspension was transferred to a sip tube for sorting. The fluorescence gate was set to sort the top 2% of fluorescent spheroplasts. The sorted volume of spheroplasts was immediately resorted with the same 2% fluorescence gate. The settings for sorting were the following: FSC/SSC= 488 nm-10 mW, SSC= 500 V Det. 2, FL1= 95/5-530/40 Det. 4= 800 V.

Rescue PCR and subcloning for g3p library

Sorted spheroplasts (10 μL of sorted volume per PCR reaction) were subjected to rescue PCR using a forward primer (BRH06= GCGGATAACAATTTCAGACAG) and a reverse primer (VJD200= CGCCATTTTTCACTTCACAGGTC) for the randomized insert. After heating the spheroplasts for 7 minutes at 95° C, 25 cycles of heating to 94° C for 30 seconds, 55° C for 30 seconds, and 72° C for 1 minute were performed followed by a final heating to 72° C for 7 minutes. Both the pAK200 vector and the PCR products were double digested simultaneously with XbaI and HindIII for 2 hours at 37° C and purified by 1% agarose gel electrophoresis. The digested PCR products and vector were ligated using T4 ligase for 24 hours at 17° C and purified by phenol-chloroform extraction and desalted by ethanol precipitation. The subcloned vector was transformed into Jude 1 cells by electroporation.

Population and colony scanning

After each round of sorting and subcloning, the population of spheroplasts was scanned by FACS for comparision to the unsorted population. After a second round of sorting and subcloning, 10 individual colonies were picked from the agar plates and used to inoculate 3 mL TB with 2% glucose supplemented with chloramphenicol. The culture was shaken overnight at 30° C. A 30 μL aliquot was subcultured into an additional 3 mL of TB with 2% glucose and chloramphenicol. The subculture was shaken for 2 hours at 37° C. The culture was induced and spheroplasts were formed as described above. Labeling was performed as described above.

4.7.7 Epitope Constructs

The expression, spheroplast formation, labeling and scanning for binding for all epitope constructs was performed in a fashion similar to the colony scanning procedure described above except the overnight culture was inoculated from a cell stock.

Streptavidin binding peptide

Both the g3p and NlpA fusions of the streptavidin binding peptide (SBP) were constructed using two PAGE purified, 5' phosphorylated, complementary oligonucleotides containing the SBP sequences. The oligonucleotides for the g3p fusion were designed to complement the sticky ends of the SfiI restriction sites of the pAK200 vector at both the 5' and 3' ends. For the pAK200 vector, the forward primer (N12-1FP = CTGCTATGGCACGTCTAGAGATCTGCCAGAACGTCTGCTACTATCTTGGAGC AGCCT) and reverse primer (N12-1BP = AGGCTGCTCCAAGATAGTAGCAGAC GTTCTGGCAGATCTCTAGACGTGCCATAGCAGGCT) were annealed and ligated to the SfiI double-cut vector. For the pAPEx vector, the forward primer (N12-2FP = CTGCTTCTAGTCGTCTAGAGATCTGCCAGAACGTCTGCTACTATCTTTAGTGA TACCGCA) and the reverse primer (N12-2BP = AGCTTGCGGTATCACTAAAG

ATAGTAGCAGACGTTCTGGCAGATCTCTAGACGACTAGAAGCAGGCT) were annealed and ligated to the SfiI and HindIII digested vector. Both vectors were purfied by phenol-chloroform extraction and ethanol precipitation. The vectors were transformed by electroporation into Jude 1 cells.

FLAG eptiopes

For the NlpA-FLAG fusion, the forward primer (FLAG-1FP = CTGACTACAAGGACGACGATGATAAGTAGTGATACCGCA) and the reverse primer (FLAG-1BP = ACGTTGCGGTATCACTACTTATCATCGTCGTCCTTGTAG TCAGGCT) were purchased PAGE purified as 5' phosphorylated and annealed and ligated to the SfiI and HindIII digested vector. The vectors were purified by phenol-chloroform extraction and ethanol precipitation and transformed into Jude 1 cells. The FLAG-SBP, FLAG-FLAG-SBP and SBP-FLAG fusion vectors were all constructed in a similar fashion with the primers listed below.

FLAG-SBP fusion:

Forward primer FLAGSA-1FP = CTGACTACAAGGACGACGATGATAAGC GTCTAGAGATCTGCCAGAACGTCTGCTACTATCTTTAGTGATACCGCA

Reverse primer FLAGSA-1BP = AGCTTGCGGTATCACTAAAGATAGTAG CAGACGTTCTGGCAGATCTCTAGACGCTTATCATCGTCGTCCTTGTAGTCAGG CT

FLAG-FLAG-SBP fusion:

Forward primer FLAGSA-2FP = CTGACTACAAGGACGACGATGATAAGG ACTACAAGGACGACGATGATAAGCGTCTAGAGATCTGCCAGAACGTCTGCTA CTATCTTTAGTGATACCGCA

Reverse primer FLAGSA-2BP = AGCTTGCGGTATCACTAAAGATAGTAGC AGACGTTCTGGCAGATCTCTAGACGCTTATCATCGTCGTCCTTGTAGTCCTTA TCATCGTCGTCCTTGTAGTCAGGCT

SBP-FLAG fusion:

Forward primer SAFLAG-1FP = CTCGTCTAGAGATTGCCAGAACGTCTGC TACTATCTTGACTACAAGGACGACGATGATAAGTAGTGATACCGCA

Reverse primer SAFLAG-1BP = AGCTTGCGGTATCACTACTTATCATCGT CGTCCTTGTAGTCAAGATAGTAGCAGACGTTCTGGCAGATCTCTAGACGAGG CT

Chapter 5

Applications of Cyclic NDI and DAN Molecules

5.1 Chapter Summary

Introduction

During the synthesis of the non-natural NDI amino acid, a side product was discovered that would allow the creation of cyclic NDI dimers. While the previous chapters involved the use of NDI and DAN interactions in the context of peptides and proteins, this chapter describes the use of a cyclic NDI dimer in DNA intercalation and classic NDI:DAN aromatic donor-acceptor interactions.

Goals

It is the aim of this chapter to explore the use of cyclic NDI and DAN assemblies. The goals were to determine if the cyclic NDI assemblies (1) provide enhanced specificity and affinity in DNA intercalation over a linear equivalent and (2) enhance affinity for a DAN monomer. In addition, experiments were designed to more accurately define the association of water soluble NDI and DAN monomers using cyclic assemblies to prevent multimeric stacking.

Approach

An NDI *bis*-amino acid unit was synthesized by formation of a symmetrical diimide using the trityl diaminopropionic benzyl ester **3.15**. This *bis*-amino acid was used to form cyclic and linear NDI dimers for DNA interacalation and binding of a DAN monomer. The interaction with DNA was characterized by DNAse I footprinting,

dissociation kinetics, and NMR, while binding of a DAN monomer was measured via ITC. In addition, the association of cyclic NDI and DAN monomers was measured by NMR titration for comparison with previously published reports.

5.2 CYCLIC INTERCALATION

5.2.1 Background

While there are a variety of DNA binding motifs, threading polyintercalation has offered modularity, versatility, and, most importantly, slow dissociation kinetics. (Chu 2007a, Guelev 2001, Guelev 2002, Lee 2004) The previously published threading polyintercalators involved at least two NDI units with peptidic linkers in between. The peptidic linkers provided both the groove selectivity (major versus minor) and sequence specificity. The linker specific to the major groove, G_3K, had a composition of three glycine residues followed by a lysine residue, and preferred the double stranded sequence GGTACC. (Guelev 2001) The linker specific to the minor groove, β-Ala_3K, had a composition of three β-alanine residues followed by a lysine residue and prefers the double stranded sequences GATAAG. (Guelev 2002) The β-alanine residues allowed the linker to span the longer, deeper minor grove as opposed to the shorter, shallower major groove (Figure 5.1).

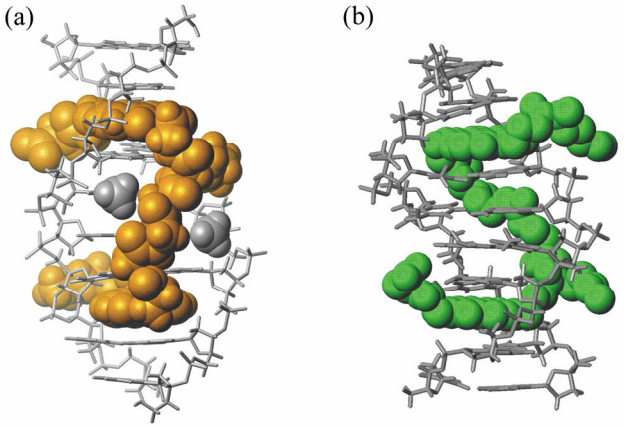


Figure 5.1 NMR structures of groove specific linkers. (a) G_3K dimer and CGGTACCG complex. (Guelev 2001) (b) β-Ala_3K dimer and CGATAAGC complex. (Guelev 2002)

In order to afford higher affinity for a specific sequence of DNA, a rigidified linker was designed. (Chu 2007a) While the rigidified bisintercalator provided greater binding affinity, the sequence specificity was diminished. An alternate design for greater binding affinity would be a sequence-specific cyclic bisintercalator. The dissociation kinetics of a cyclic binding motif would be extremely slow compared with a linear bisintercalator and sequence specificity could be enhanced by binding both grooves of a DNA sequence with two linkers simultaneously (Figure 5.2). It is unclear as to the mechanism of binding of the cyclic bisintercalator with DNA. Double stranded DNA is known to be dynamic in solution with hydrogen bonding of base pairs breaking and reforming rapidly at room temperature. A cyclic bisintercalator could take advantage of the "breathing" of duplex DNA. The "opening" and "closing" of base pairs has been demonstrated by the observation of imino proton half-lives in the 1-50 ms range. (Gueron 1992)

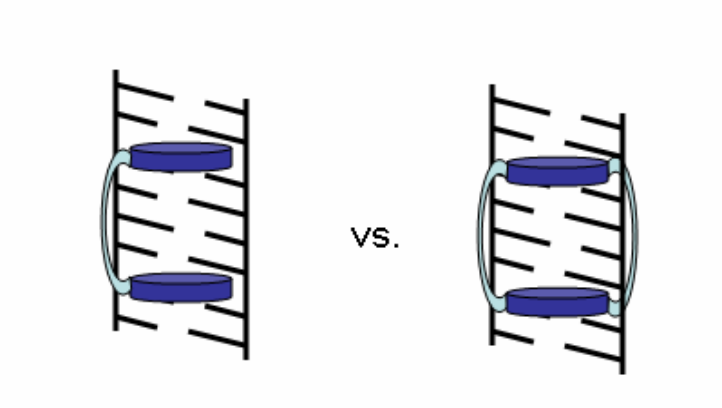
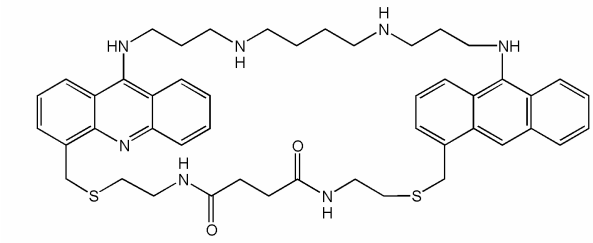


Figure 5.2 Cartoon of classical threading bisintercalation versus cyclic bisintercalation.

Cyclic DNA bisintercalators have been studied previously, albeit with little success in programming specificity. Bisacridine A (Figure 5.3a) was studied for binding to double stranded DNA, although multiple binding modes were found. The bisacridine A was found to bind the sequences CGCG by spanning either two base pairs or a single base pair. (Zimmerman 1989, Veal 1990) Bisacridine B (Figure 5.3b) was determined to bind a DNA 11-mer, where one acridine intercalated an abasic site. While this was determined to be the major complex, an additional complex was found but not structurally characterized. (Jourdan 1999)

(a)



(b)

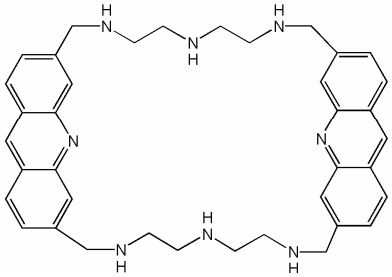


Figure 5.3 Cyclic Bisintercalators (a) bisacridine A and (b) bisacridine B.

During optimization of the synthesis of the non-natural NDI amino acid **3.18**, a major side product was discovered in the formation of asymmetrical diimide **3.16**. In addition to the correct unsymmetrical diimide, the symmetrical bis-amino acid diimide was formed (Figure 5.4). The NDI *bis*-amino acid was considered for use in creating a cyclic bisintercalator. In this section, the design and synthesis of a cyclic bisintercalator utilizing the NDI *bis*-amino acid is presented. In addition, the results of experiments to characterize the cyclic bisintercalator-DNA interaction by dissociation kinetic behavior with poly (dGdC), sequence specificity by DNAse I footprining, and NMR will be discussed.

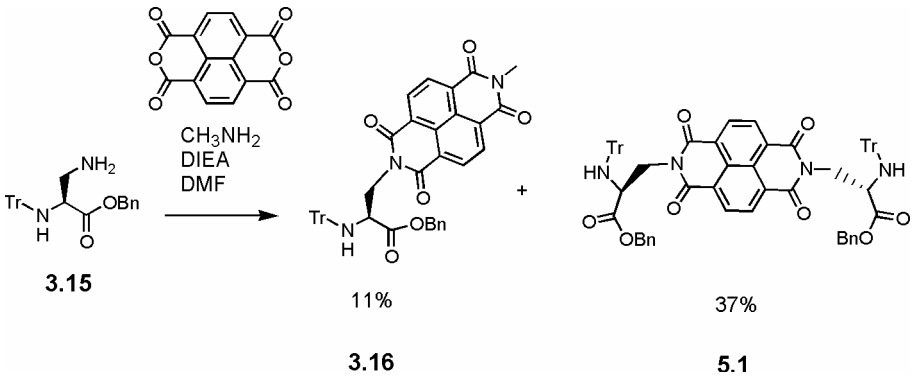


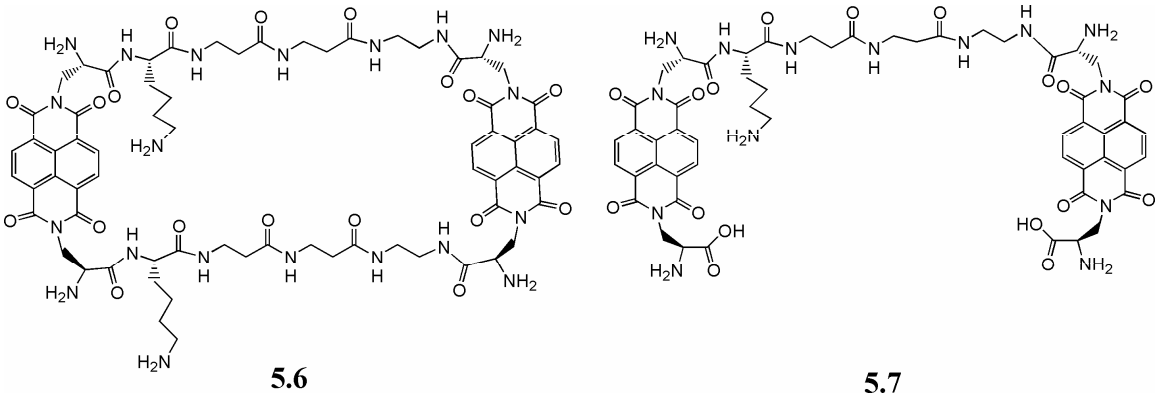
Figure 5.4 Diimide formation of NDI non-natural amino acid **3.16**, where the NDI *bis*-amino acid **5.1** was a major side product.

5.2.2 Results and Discussion

5.2.2.1 Cyclic bBisintercalator Design

The first generation of cyclic bisintercalator was designed with two minor groove β-Ala_3K linkers in a symmetrical fashion as shown in Figure 5.5. While the linker is only specific for the minor groove and not the major groove of the GATAAG sequence, it was chosen for its length. Since minor groove binding is a longer path than the major groove, use of the shorter, major groove G_3K linker would not provide the length required for

spanning both grooves and could have deleterious consequences for binding. Also, it was hypothesized that the β-Ala_3K linker would provide sequence specificity in the minor groove even though binding of the same linker in the major groove may not be favorable. To accommodate the NDI *bis*-amino acid in the cyclic structure, the peptide linkers needed two amino termini. This was accomplished by replacing what would be the *C*-terminal β-alanine residue with an isostructural ethylendiamine. In addition to the cyclic bisintercalator, an analogous linear bisintercalator was also designed for comparison.

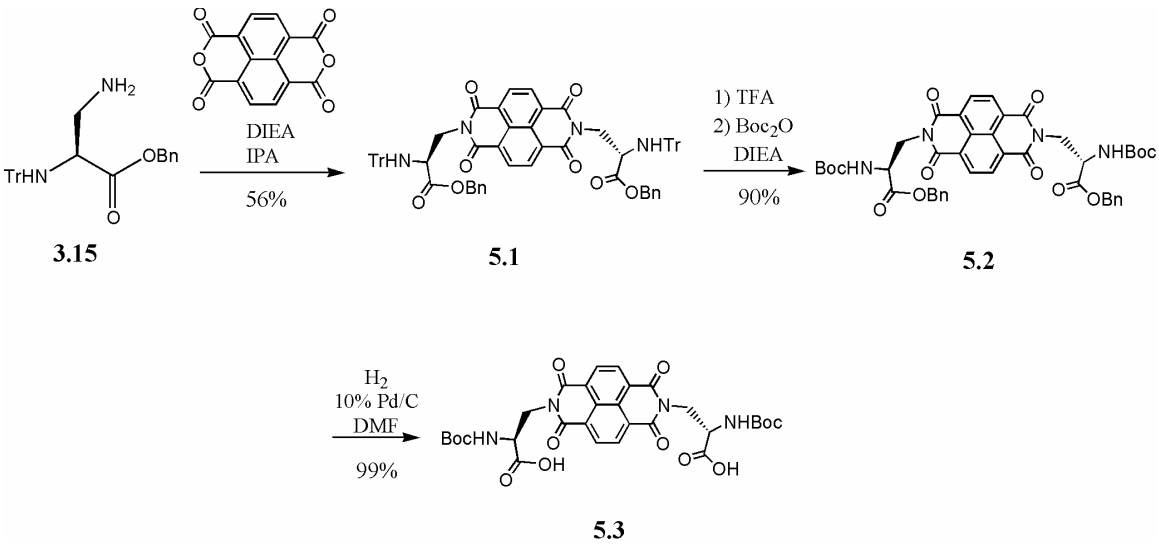


5.6 **5.7**

Figure 5.5 Design of a cyclic bisintercalator **5.6** and the linear analogue **5.7**.

5.2.2.2 Cyclic/Linear Bisintercalator Synthesis

The cyclic and linear bisintercalators were synthesized according to Scheme 5.1 and 5.2. The NDI *bis*-trityl amino benzyl ester **5.1** was synthesized by refluxing trityl diaminopropionic benzyl ester **3.16** with 1,4,5,8-naphthalenetetracarboxylic dianhydride. Compound **5.1** was trityl deprotected and reprotected with a Boc group to afford compound **5.2**. The benzyl esters were hydrogenated to afford the NDI *bis*-Boc amino acid **5.3**.

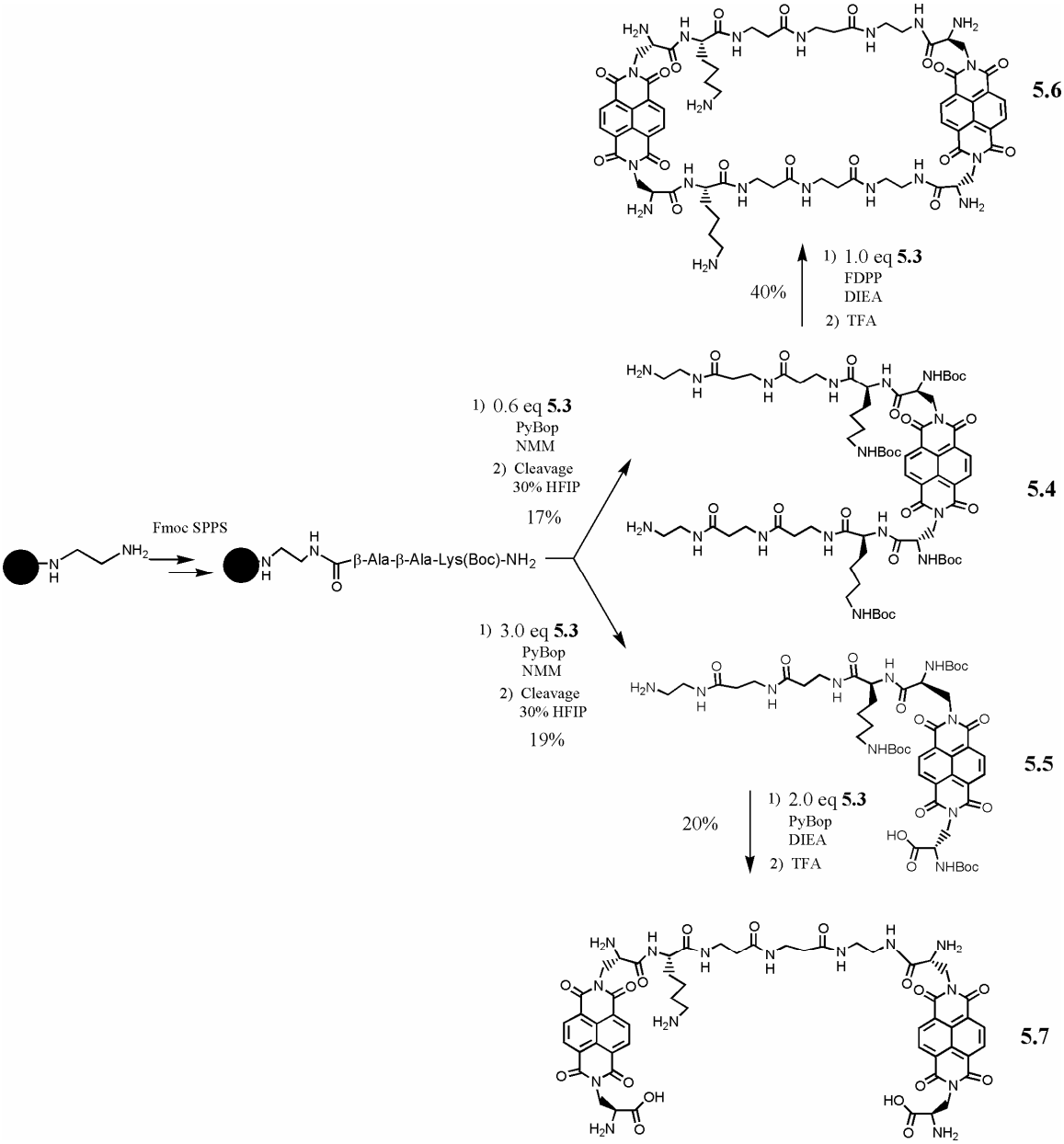


3.15 **5.1** **5.2** **5.3**

Scheme 5.1 Synthesis of the Boc protected NDI bis-amino acid **5.3**.

The β-Ala$_3$K linker was synthesized on pre-loaded ethylenediamine-2-chlorotrityl resin. The 2-chlorotrityl resin was chosen for its mild cleavage conditions. Amines are easily cleaved from 2-chlorotrityl chloride resin with fluorinated reagents such as trifluoroethanol and hexafluorisopropanol, without Boc deprotection of the side chains. Coupling compound **5.3** to the resin followed by cleavage with hexafluoroisopropanol provided both the cross-linked monomer **5.4** for the cyclic dimer as the major product and single coupled monomer **5.5** for the linear dimer, which were seperable by RP-HPLC. Alternatively, coupling of compound **5.3** was optimized to produce the single coupled monomer **5.5** as the major product. The cyclic dimer **5.6** was synthesized by coupling limiting amounts of compound **5.3** with compound **5.4** under very dilute conditions, followed by global Boc deprotection with TFA. Interestingly, the retention of compound **5.6** on RP-HPLC was almost identical to the precursor compound **5.4**. The hydrophilicity of the additional four positive charges is nearly cancelled by the increase in hydrophobicity of an additional NDI unit. The linear dimer **5.7** was synthesized by

coupling excess amounts of coumpound **5.3** with compound **5.5** under very dilute conditions followed by global Boc deprotection with TFA.



Scheme 5.2 Synthesis of cyclic bisintercalator **5.6** and linear bisintercalator **5.7**.

5.2.2.3 Dissociation Kinetics

Previous work has established that NDI bisintercalators bind poly(dGdC) DNA more tightly than poly(dAdT), therefore, dissociation kinetic experiments were performed with poly(dGdC) as an initial test of binding. (Chu 2007b) Binding of **5.6** with poly(dGdC) was indicated by the characteristic hypochromism in the NDI absorbance at 363 and 386 nm compared to the free bisintercalator as shown in Figure 5.6a and b.

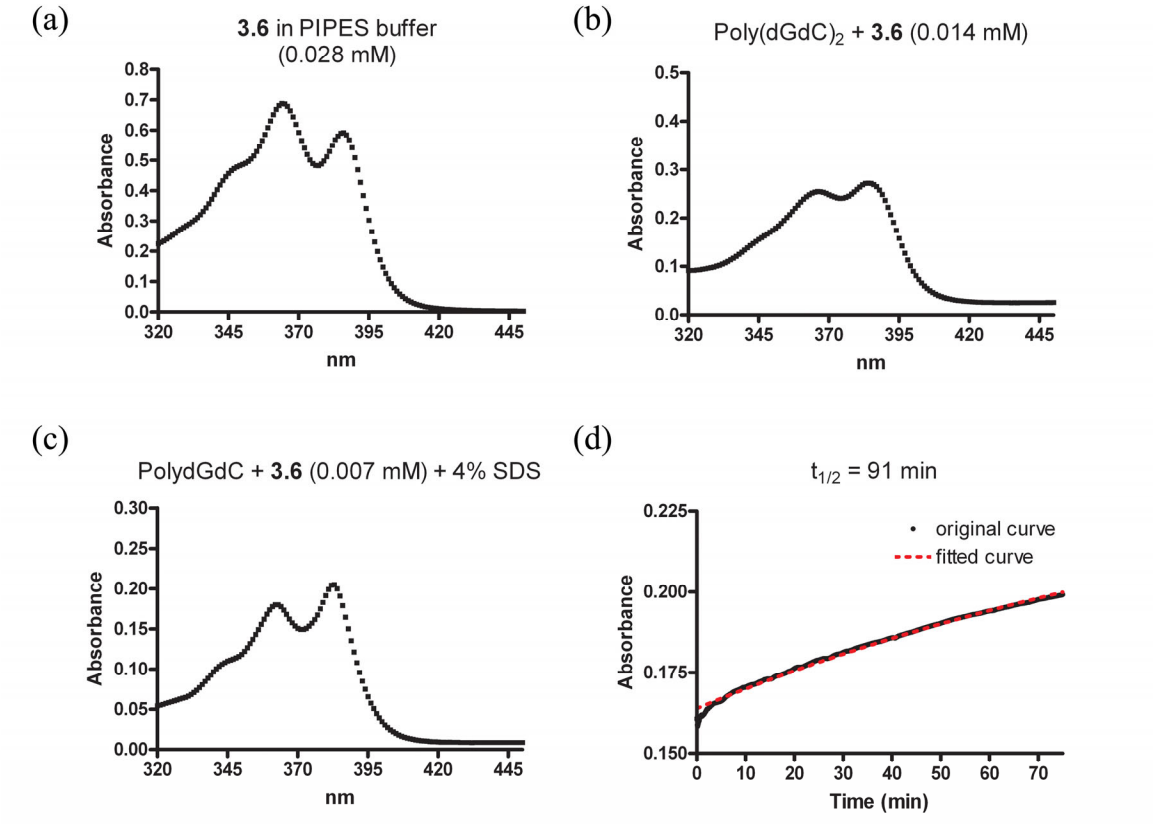


Figure 5.6 Absorbance spectra of (a) **5.6** in buffer, (b) **5.6**-poly(dGdC) complex, (c) **5.6**-poly(dGdC) with 4% SDS, and (d) the dissociation kinetics of the **5.6**-poly(dGdC) complex with 4% SDS.

The dissociation kinetics in Figure 5.6c was fit to a one phase exponential decay, $A = a_0e^{(-kt)} + b_0$ (k = dissociation rate constant). (Chen 1997) The half-life of **5.6** with

poly(dGdC) was determined to be 91 minutes. The typical linear dimer had a half-life of 15 minutes, while another version of the cyclic bisintercalator had a half-life of 41 minutes (Figure 5.7). (Chu 2007b) The dissociation kinetics of **5.6** appeared to be almost linear and did not fit the one phase exponential decay well. This may be due to multiple binding modes or non-specific binding of the positively charged **5.6** to the highly negatively charged DNA.

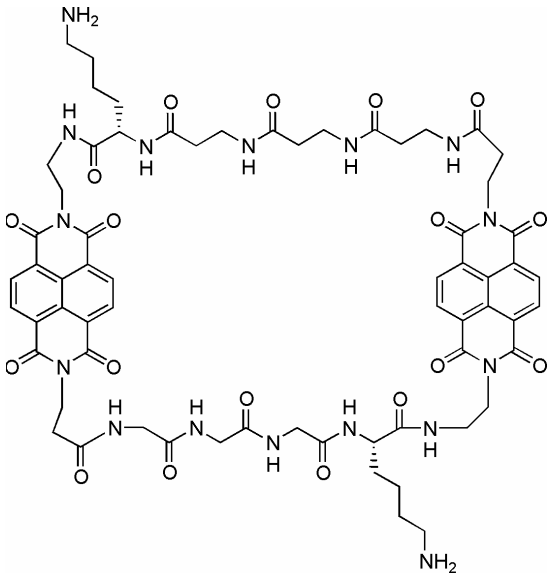


Figure 5.7 Structure of an alternate NDI-based cyclic bisintercalator that incorporated both the G_3K and β-Ala_3K linkers. (Chu 2007b)

5.2.2.4 DNAse I Footprinting

While the binding site for the cyclic intercalator was designed to be GATAAG for the β-Ala_3K linker, DNAse I footprinting was performed to determine the preferred binding sequences. A 69-mer of DNA (Figure 5.8a) was used that incorporated multiple known binding sites of NDI-containing threading polyintercalators including GGTACC, GATAAG, CTTACT, and GATAAG.

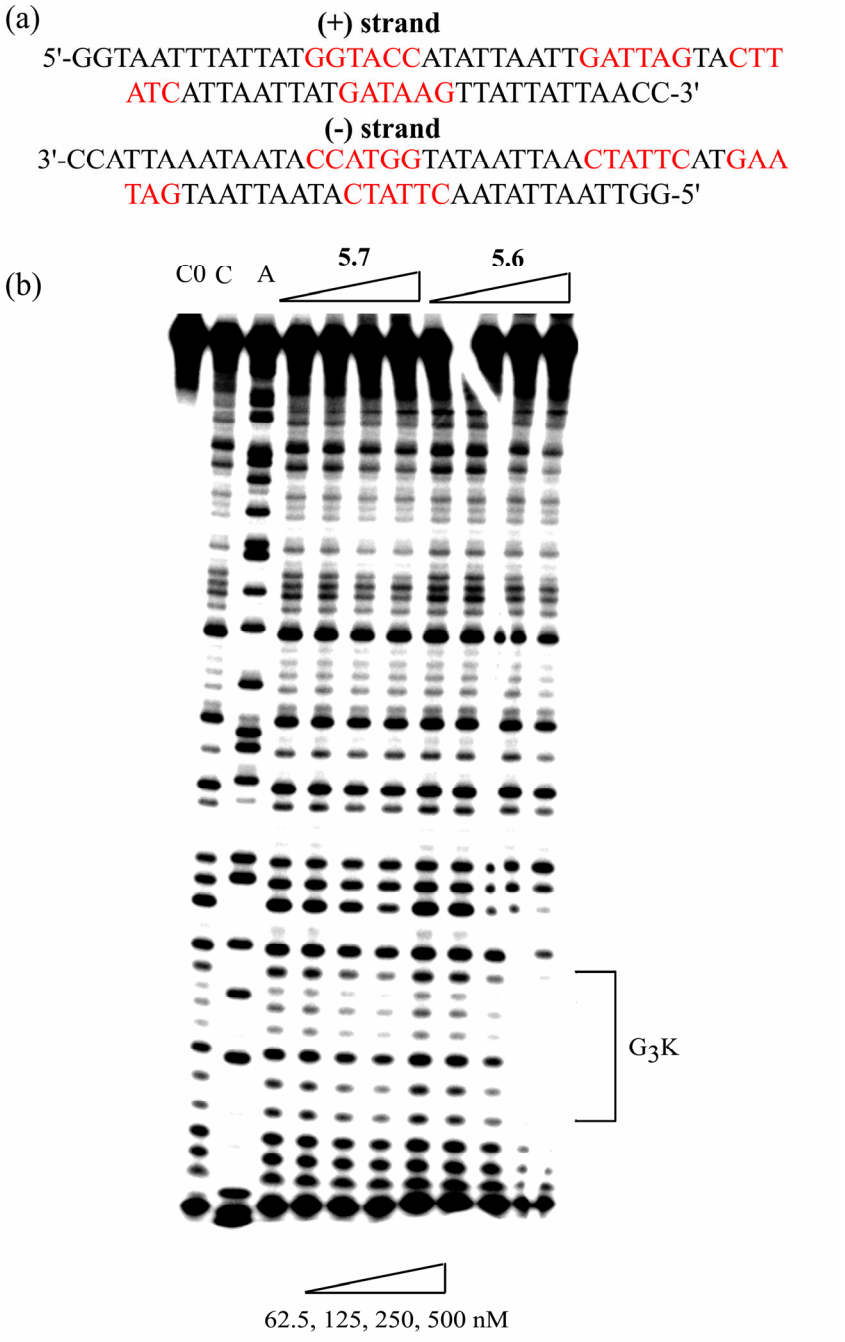


Figure 5.8 (a) The 69-mer DNA sequence used for DNAse I footprinting. (b) The DNAse I footprinting of the cyclic bisintercalator **5.6** compared to the linear bisintercalator **5.7**. Lane A represents an adenine-specific sequencing reaction. Lane Co represents DNA without reaction with DNAse I. Lane C represents DNA with DNAse I and without ligand. G_3K sequence GGTACC is highlighted.

The DNAse I footprinting data is shown in Figure 5.8b. Unexpectedly, the cyclic bisintercalator did not show a footprint for the β-Ala_3K sequence GATAAG. Instead, there was a footprint for the sequence GGTACC, which was the sequence for the G_3K linker. The corresponding linear dimer **5.7** did not show a well defined footprint at 500 nM concentration. Interestingly, the CBI-1 cyclic bisintercalator in Figure 5.7 also displayed affinity for the GGTACC sequence. Since CBI-1 contains both the β-Ala_3K and G_3K linkers, this affinity was thought to arise from the domination of the G_3K recognition in the major groove. These results, however, suggest the sequence specificity may arise from the general design of this cyclic structure rather than the linker sequence. Additional work needs to be done, including synthesis of CBI-1 with two β-Ala_3K linkers and **5.6** with two G_3K linkers, to fully understand the factors underlying the sequence specificity of these cyclic bisintercalators.

5.2.2.5 NMR Analysis

A titration of d(CGGTACCG) with **5.6** was monitored by 1D ^{1}H NMR spectroscopy. Upon addition of one equivalent of **5.6** to the free DNA for 2 hours, no significant change in position of the imino protons of G_2, G_3 and T_4 was seen. To determine if the lack of binding of **5.6** to this sequence was a result of slow kinetics of binding, the sample was heated to 50° C for 1 hour then cooled to 27° C for analysis. The peaks became very broad with no significant shifting of the imino protons except very small, broad peaks near 11 ppm as shown in Figure 5.9. The lack of sharp, upfield shifted peaks may be due to multiple binding modes with this short sequence or non-specific binding without intercalaction. It is unclear why the footprinting showed sequence specificity and the UV spectrum with poly(dGdC) showed the hypochromism signature of intercalation, while the NMR does not show a specific binding mode. The length of the sequence for NMR analysis may play a role in this difference. The NMR analysis,

however, confirmed the dissociation kinetic studies that did not conform to a one phase exponential decay as a result of multiple binding modes.

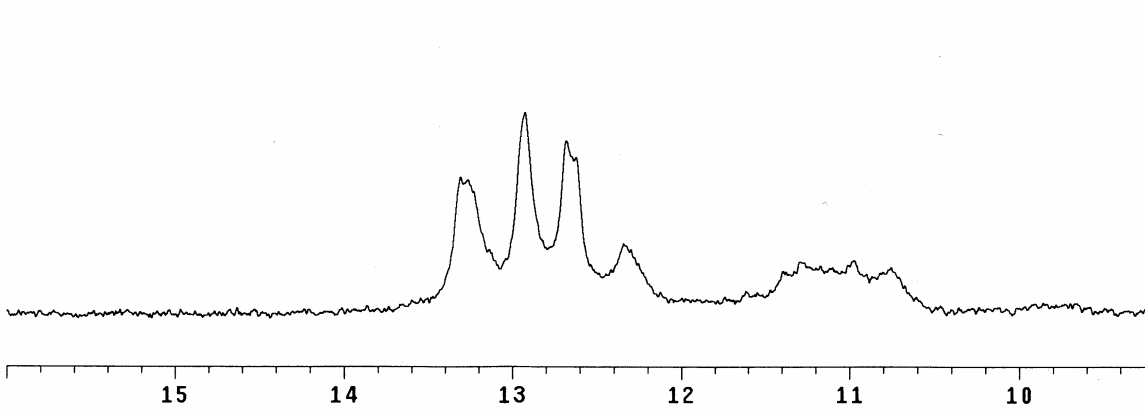


Figure 5.9 1D ^{1}H NMR experiment of a 1:1 complex of **5.6** and CGGTACCG after heating to 50° C in the G_2, G_3 and T_4 imino proton region.

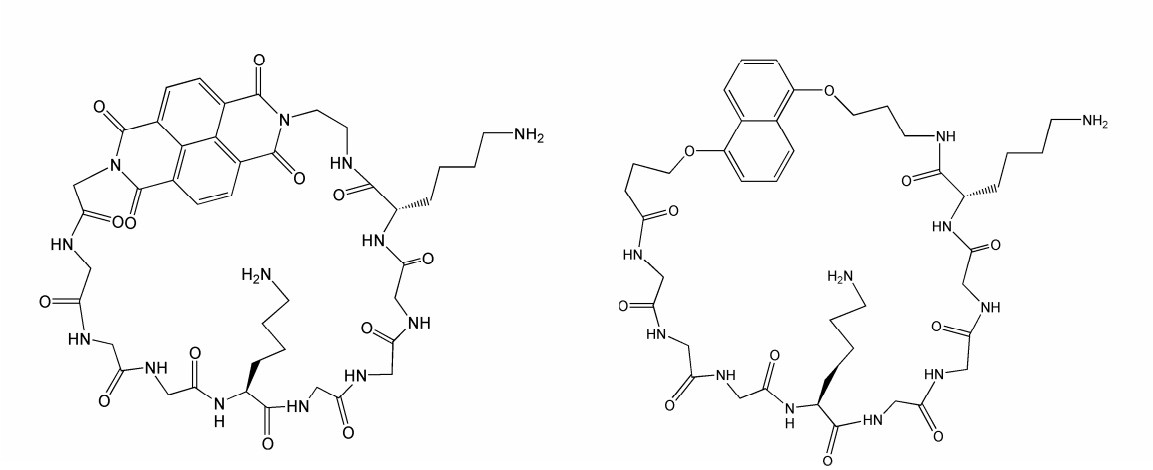
5.3 CYCLIC ASSEMBLIES FOR AROMATIC DONOR-ACCEPTOR INTERACTIONS

5.3.1 Background

While the cyclic and linear NDI dimers showed little promise for use in DNA intercalation, they had utility in probing NDI-DAN interactions. Previous work described in Chapter 1 had shown the interaction of monomers, dimers, and higher oligomers of negatively charged NDI and DAN units in an intermolecular assembly. (Gabriel 2002) It was calculated that replacing the charged system with a neutral one could potentially increase the association constant by an order of magnitude per interaction. However, neutral DAN oligomers were not sufficiently soluble to study the hypothesis extensively. Limited data suggested a tentative $\Delta G°$ of -10.2 kcal/mol for a neutral DAN tetramer with a -4 charged NDI tetramer, which would imply an increase of two orders of magnitude in association over the charged DAN equivalent. (Reczek 2006) While an intramolecular assembly of a non-alternating DAN-DAN-NDI folded trimer structure was supported by 2D-NMR analysis, the intermolecular interaction of a charged

NDI dimer with a neutral DAN monomer has not been probed experimentally. (Gabriel 2005) The cyclic NDI dimer could provide insight into how a conformationally restricted cyclic NDI dimer would affect binding of a neutral DAN monomer as compared to the less restricted linear NDI dimer counterpart.

Cyclic assemblies could also provide conformation of previous measurements of NDI-DAN interactions. Previous work with neutral monomers showed a NDI-DAN binding constant of $2x10^3$ M^{-1} in D_2O using a model that accounted for the multiple equilibria in the Scatchard plot. (Cubberly 2001) By using cyclic, positively charged NDI and DAN monomers (Figure 5.10) synthesized in a manner analogous to CBI-1, the multimeric stacking of the monomers could be potentially eliminated. In addition, the effect of charge type (positive versus negative) could be probed by comparison to the negatively charged monomer that displayed an association constant of 200 M^{-1}. (Gabriel 2002)



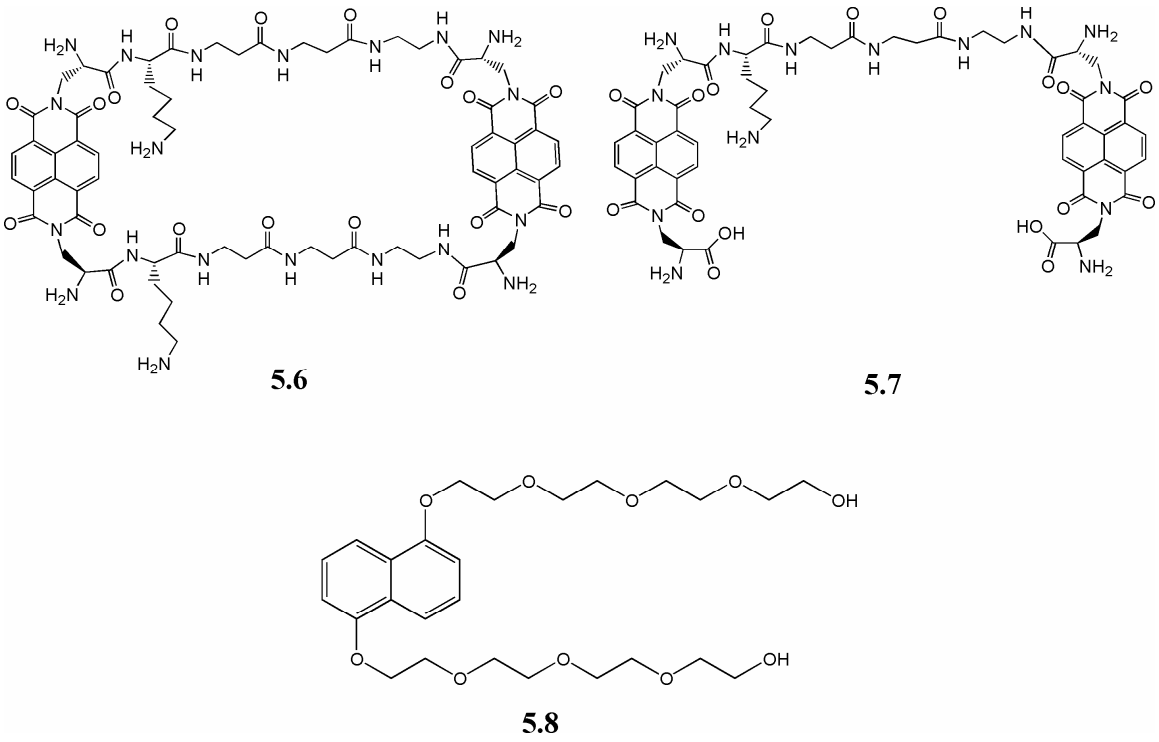
5.9 **5.10**

Figure 5.10 Structures of positively charged cyclic NDI (**5.9**) and DAN (**5.10**) monomers.

5.3.2 Results and Disscussion

5.3.2.1 Isothermal Titration Calorimetry of Cyclic and Linear NDI Dimers

In order to determine how a conformationally restricted cyclic NDI dimer would affect binding of a neutral DAN monomer as compared to the less restricted linear NDI dimer counterpart, isothermal calorimetry titrations of the cyclic (**5.6**) and linear (**5.7**) NDI dimers with the neutral DAN monomer **5.8** (Figure 5.11) were performed in degassed 50 mM sodium phosphate buffer at pH 7. To avoid dilution effects, the neutral DAN monomer was titrated into the cell containing the charged NDI dimers.



5.6 **5.7**

5.8

Figure 5.11 Structures of the positively charged cyclic (**5.6**) and linear (**5.7**) NDI dimers and the neutral DAN monomer **5.8**.

Titration of the cyclic NDI dimer **5.6** with the neutral DAN monomer gave an association constant of 223 M^{-1} with a $\Delta H°$ value of -6 kcal/mol and a $\Delta S°$ value of -9.52

cal·mol^{-1}·K^{-1} as shown in Figure 5.12a. Titration of the linear NDI dimer **5.7** with the neutral DAN monomer gave an association constant of 725 M^{-1} with a ΔH° value of -2.9 kcal/mol and a ΔS° value of 3.15 cal·mol^{-1}·K^{-1} as shown in Figure 5.12b.

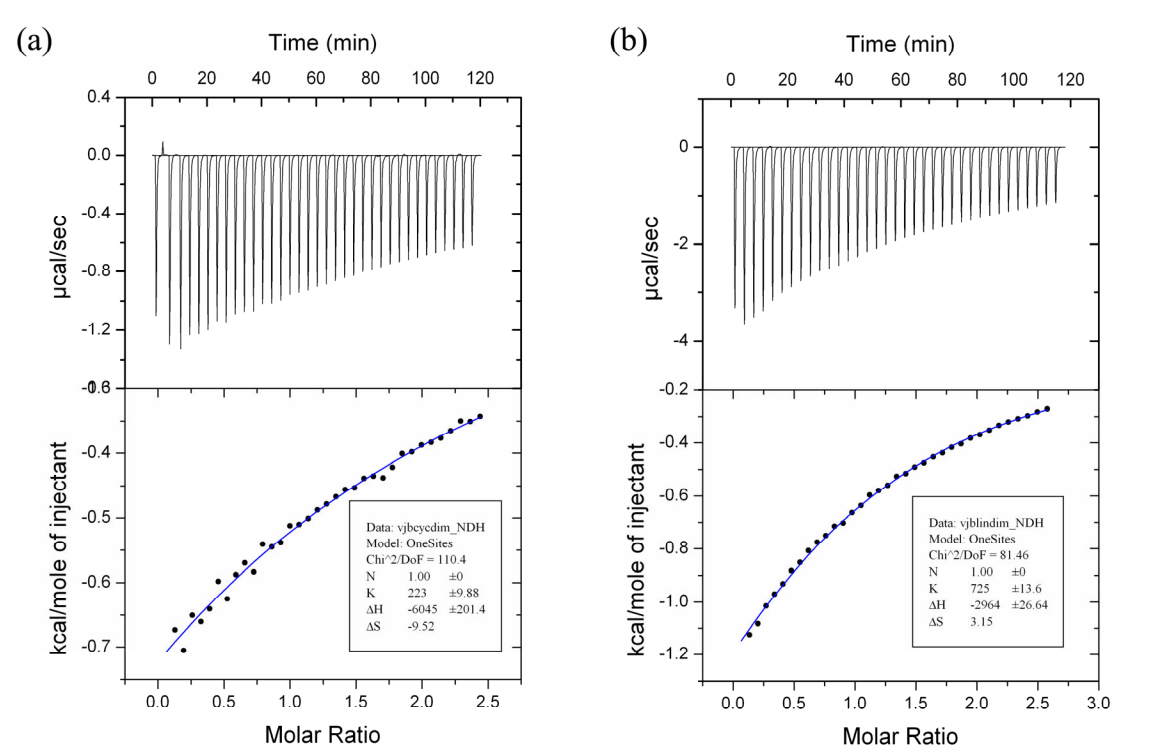


Figure 5.12 Isothermal calorimetry titrations of neutral DAN monomer **5.8** into (a) 0.6 mM cyclic NDI dimer **5.6** and (b) 0.9 mM linear NDI dimer **5.7** in 50 mM sodium phosphate buffer at pH 7.

Interestingly, binding of the neutral DAN monomer to the cyclic NDI dimer **5.6** gave an association constant three times less than the linear NDI dimer **5.7**. This was not completely unexpected as the cyclic dimer was designed for DNA intercalation and the linkers were longer than required for complexation of the DAN monomer. This was obvious from the entropy term of the cyclic dimer complex. The negative entropy of the

cyclic dimer is a result of severe restriction of the degrees of freedom of the long, already constrained linkers of the cyclic dimer complex in order to achieve binding. In addition, the same argument can be used to explain the larger enthalpic driving force for binding to the cyclic dimer. The long, constrained linkers of the cyclic structure may prevent complete self-stacking of the NDI units in the initial structure making the binding of the DAN monomer and desolvation of the NDI faces enthalpically more favorable. The linear dimer, however, had half the enthalpic driving force because the single linker does not provide enough conformational restriction to prevent self-stacking of the NDI units. Since the starting structure already hides the faces of the NDI units, switching the NDI self-stacking for the NDI-DAN interaction is not nearly as downhill enthalpically as for the cyclic dimer. By the same token, the difference in structure between the starting self-stacked linear dimer and the linear dimer-DAN complex is not large; therefore, the lack of any additional conformational restriction does not require an additional entropic cost.

5.3.2.2 NMR Titration of Cyclic Monomers

The cyclic NDI monomer **5.9** was titrated at 0.375 mM concentration in 50 mM sodium phosphate at pH 7 with 0 to 8 equivalents of cyclic DAN monomer **5.10**. The data was fitted to two different equations (Figure 5.13):

1-to-1 Wilcox: $\delta = \delta_{ini} + (\Delta\delta_{tot}/[H])[(1/K +[H]+[G])-\sqrt{((1/K +[H]+[G])^2-4[H][G])}]$

1-to-1 Connors: $\delta = (\Delta\delta_{tot}K[G])/(1+K[G])$

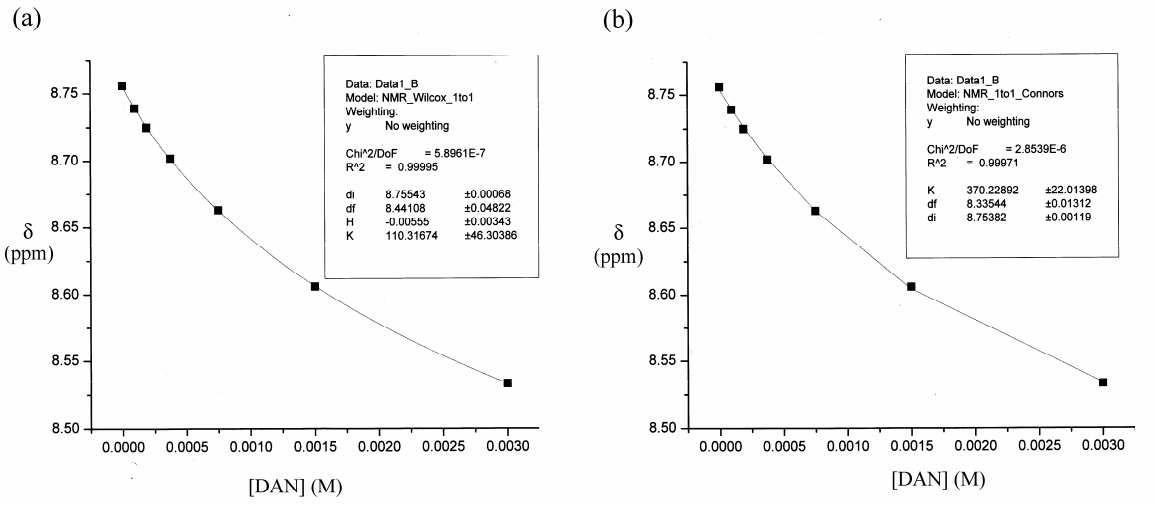


Figure 5.13 Comparison of the NMR titration experiment of 0.375 mM cyclic NDI monomer 5.9 with 0 to 8 equivalents of cyclic DAN monomer 5.10 fitted with (a) the 1-to-1 Wilcox and (b) 1-to-1 Connors equations.

The resulting association constant of the Wilcox equation was 110 M^{-1}. However the concentration of the cyclic NDI monomer was determined to be negative with this equation. (Wilcox 1991) There are two explanations for this result, the first of which is that there were not enough data points for the Wilcox equation to be used accurately and the concentrations were not high enough. Additionally, this may suggest multimeric stacking had not been excluded by the use of cyclic monomers since even after 8 equivalents of cyclic DAN monomer, the plot had not leveled off. The resulting association constant from the Connors equation was 370 M^{-1}. (Connors 1987) This was consistent with the 200 M^{-1} association constant determined for the negatively charged monomers.

5.4 CONCLUSIONS

The cyclic bisintercalator **5.6** was designed and synthesized in addition to the linear counterpart **5.7**. While DNAse I footprinting showed sequence specificity for the

GGTACC sequence and hypochromism demonstrated intercalation, dissociation kinetics displayed near linear behavior most likely from a combination of intercalation and non-specific association of the +6 charged bisintercalator with the highly negatively charged DNA. NMR analysis of the footprint site with compound **5.6** indicated multiple binding modes of very little intercalated compound. While the side chains of the NDI *bis*-amino acid have the same number of atoms as the NDI unit of the alternative cyclic intercalator in Figure 5.7, the amino group of the *bis*-amino acid may be providing steric bulk and preventing full intercalation. It would be interesting to make a similar cyclic bisintercalator using the longer diaminobutyric acid (Dab) residue from Chapter 3 for the NDI *bis*-amino acid as a comparison. In addition, the origin of the specificity of the cyclic bisintercalators for the G_3K sequence should be probed.

Isothermal titration calorimetry experiments with the cyclic and linear NDI dimers revealed a lowered association constant for the cyclic dimer due to a severe entropic penalty for binding of a neutral DAN monomer. To optimize the binding of the cyclic versus linear dimers, modeling needs to be performed in order to determine the best length of the linkers for DAN binding. In addition, the linkers should be rigid to prevent self-stacking of the NDI units to maximize the enthalpic gain of the NDI-DAN interaction. Work is currently underway in the group to synthesize a rigid, water soluble linker with the correct spacing.

NMR titration of cyclic NDI and DAN monomers confirmed previous results of the association constant between charged monomers (370 M^{-1} versus 200 M^{-1}). Additional compound would be required to repeat the measurements at higher concentration and more accurately measure the association constant.

5.5 EXPERIMENTAL METHODS

General Procedures Resins, amino acids, and coupling reagents were purchased from Novabiochem. All other starting materials were purchased from Aldrich and used without further purification. Flash chromatography was performed on Silicycle Silia-P silica gel. 1H and 13C NMR spectra were recorded on a Varian Unity Plus 400 MHz spectrometer in the indicated solvent. Chemical shifts are expressed in parts per million (ppm, δ) relative to tetramethylsilane (TMS, δ = 0.00 ppm). Analytical and preparative HPLC was performed on Vydac C18 protein and peptide columns with monitoring at 384 nm, where solvent A was water containing 0.1% TFA and solvent B was acetonitrile containing 0.1% TFA. UV spectra were recorded using a Perkin-Elmer Lambda 35 spectrophotometer. Isothermal calorimetry titrations were performed on a VP-ITC microcalorimeter from Microcal Inc. (Northhampton, MA 01060). The neutral DAN monomer **5.8** was obtained from Joseph Reczek of the Iverson group. (Reczek 2006) The cyclic NDI and DAN monomers, **5.9** and **5.10**, were obtained from Yongjun Chu of the Iverson group. (Chu 2007b) Dissociation kinetics, DNAse I footprinting, and NMR analysis were performed according to previously published procedures. (Chu 2007b) Data for ITC and NMR titrations were analyzed and fitted using Origin 7.0 software (Northhampton, MA 01060).

(2S,2'S)-benzyl 3,3'-(1,3,6,8-tetraoxobenzo[*lmn*][3,8]phenanthroline-2,7(1*H*, 3*H*, 6*H*, 8*H*)-diyl)bis(2-tritylamino)propanoate) (5.1) Compound **3.15** (5.48 g, 12.6 mmol) was dissolved in isopropanol (90 ml) with diisopropylethylamine (3.84 ml, 22 mmol). 1,4,5,8-Naphthalenetetracarboxylic dianhydride (1.69 g, 6.3 mmol) was added and the reaction was refluxed overnight under a nitrogen atmosphere. The mixture was cooled to room temperature and concentrated under reduced pressure. The crude product was purified by

column chromatography (70:25:5 DCM/hexanes/ACN) to afford compound **5.1** as a yellow solid (3.91 g, 56%). ^{1}H NMR ($CDCl_3$) δ 8.62 (s, 4H), 7.52 (m, 10H), 7.32-7.15 (m, 20H), 6.88 (m, 10H), 4.70-4.55 (m, 4H), 4.50 (d, *J* = 12 Hz, 2H), 4.28 (d, *J* = 12 Hz, 2H), 3.93 (m, 2H), 3.01 (d, J= 9.6 Hz, 2H). ^{13}C NMR ($CDCl_3$) δ 172.46, 162.59, 145.47, 134.59, 130.82, 128.64, 127.87, 127.79, 126.45, 126.31, 126.11, 70.98, 69.97, 54.51, 43.74, 30.86. HRMS (ESI) calculated for $C_{72}H_{56}N_4O_8Na$ $(M+Na)^+$: m/z 1127.399, found 127.402.

(2S,2'S)-benzyl 3,3'-(1,3,6,8-tetraoxobenzo[*lmn*][3,8]phenanthroline-2,7(1*H*, 3*H*, 6*H*, 8*H*)-diyl)bis(2-(*tert*-butoxycarbonylamino)propanoate) (5.2) Compound **5.1** (2.08 g, 1.9 mmol) was dissolved in a mixture of TFA (0.87 ml, 11.3 mmol), methanol (1.1 ml), and DCM (110 ml). The reaction was stirred for 10 minutes at room temperature and then concentrated under reduced pressure. The residue was suspended in DMF (19 ml), which dissolved upon addition of diisopropylethylamine (1.97 ml, 11.3 mmol). To the solution, di-tert-butyl dicarbonate (0.90 g, 4.1 mmol) was added and the reaction was stirred at room temperature overnight under an atmosphere of nitrogen. The reaction was poured into DCM (100 ml) and washed with 0.5 M citric acid (3 x 100 ml), saturated $NaHCO_3$ (3 x 100 ml), and water (100 ml). The organic layer was dried over Na_2SO_4 and concentrated under reduced pressure. Purification of the crude product by column chromatography (5:4:1 DCM/hexanes/acetone) gave compound **5.2** as a pale yellow solid (1.40 g, 90%). ^{1}H NMR ($CDCl_3$) δ 8.69 (s, 4H), 7.28-7.21 (m, 10H), 5.40 (d, *J* = 8.4 Hz, 2H), 5.18 (s, 4H), 4.93 (q, *J* = 7.2 Hz, 2H), 4.57 (m, 4H), 1.20 (s, 18H). ^{13}C NMR ($CDCl_3$) δ 170.06, 162.95, 155.25, 134.82, 131.04, 128.40, 128.33, 128.17, 126.58, 126.26, 80.03, 67.67, 51.86, 41.83, 27.98. HRMS (ESI) calculated for $C_{44}H_{44}N_4O_{12}Na$ $(M+Na)^+$: m/z 843.285, found 843.287.

(2S,2'S)-3,3'-(1,3,6,8-tetraoxobenzo[*lmn*][3,8]phenanthroline-2,7(1*H*, 3*H*, 6*H*, 8*H*)-diyl)bis(2-(*tert*-butoxycarbonylamino)propanoic acid) (5.3) Compound **5.2** (1.94 g, 2.4 mmol) was dissolved in DMF (47 ml). 10% Palladium on carbon (0.25 g, 10 mol %) was added and the solution was purged with hydrogen gas. The reaction was stirred for 6 hours under a balloon of hydrogen. The mixture was filtered and the palladium was rinsed with DMF. The combined filtrates were concentrated under reduced pressure. The residue was dissolved in DCM, filtered to remove the remaining palladium, and the product was precipitated with hexanes. Filtration of the precipitate gave compound **5.3** as a tan solid (1.50 g, 99%). ^{1}H NMR (DMSO) δ 8.65 (s, 4H), 7.19 (bs, 2H), 4.41 (m, 6H), 1.14 (s, 18H). ^{13}C NMR (DMSO) δ 172.24, 163.28, 156.01, 131.24, 126.76, 126.71, 78.86, 51.44, 41.55, 28.52, 28.44, 28.42. HRMS (ESI) calculated for $C_{30}H_{32}N_4O_{12}Na$ $(M+Na)^+$: m/z 663.191, found 663.192.

Fmoc solid phase peptide synthesis Solid phase synthesis of the β-Ala_3K linker was performed similarly for both the cyclic **5.4** and linear dimer **5.5** precursor. Solid phase peptide syntheses were performed on pre-loaded ethylenediamine-2-chlorotrityl resin (200-400 mesh) with a loading of 1.5 mmol/g on a 0.5 mmol scale. Fmoc-β-Ala-OH (3 eq), PyBop (3 eq), and N-methyl morpholine (6 eq) were dissolved in NMP. The solution was added to the resin and shaken for 1.5 hours. The solution was drained and the resin was washed three times with DMF, IPA, and again with DMF. Deprotection of the Fmoc protecting group was afforded by shaking the resin with 20% piperidine in DMF for 20 minutes. The resin was washed again three times with DMF, IPA, and DMF. Deprotection/coupling cycles were performed to synthesize the remaining sequence on the resin.

(H_2N-β-$Ala_3K(Boc))_2$-$(Boc)_2NDI$ cross-linked monomer (5.4) Upon completion of β-Ala_3K linker synthesis on solid phase, the terminal Fmoc was deprotected with 20% piperdine. The β-Ala_3K linkers were crosslinked by coupling compound **5.3** (0.3 eq) with the resin using PyBop (0.6 eq) and N-methylmorpholine (0.6 eq) in NMP. The mixture was shaken for 1.5 hours and drained. The crosslinking was repeated with fresh reagents for an additional 4 hours. After draining, the resin was washed three times each with DMF, IPA, and DMF. The resin was cleaved using 30% hexafluoroisopropanol (HFIP) in DCM for 2 hours. The fully protected peptide was purified by preparative HPLC using a gradient of 33-35% B over 30 minutes. The product eluted at 13 minutes and lyophilization of pure fractions gave compound **5.4** as a white powder (60.7 mg, 17%). HRMS (ESI) calculated for $C_{68}H_{105}N_{16}O_{20}$ $(M+H)^+$: m/z 1465.769 , found 1465.769.

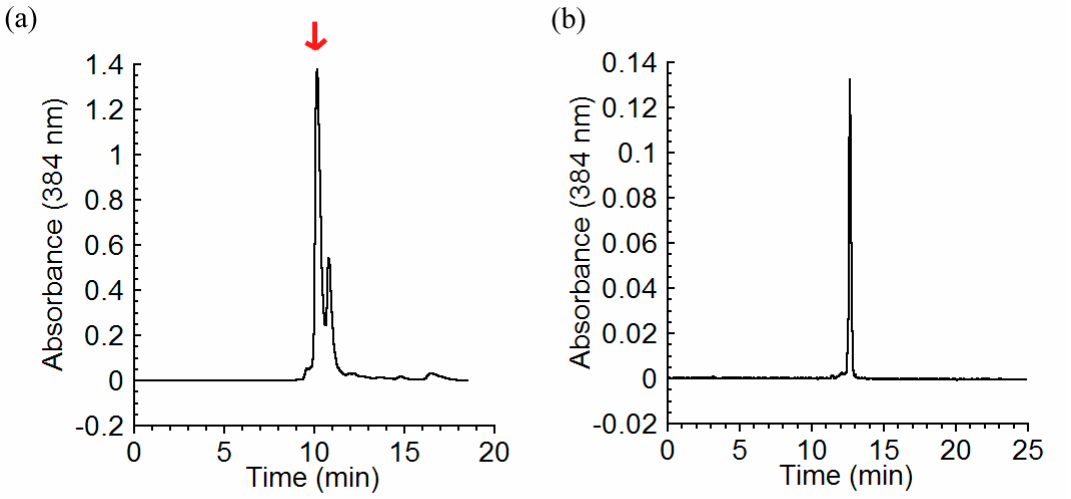


Figure 5.14 Compound **5.4** (a) crude and (b) pure HPLC. Red arrow indicates product in crude mixture.

H_2N-β-Ala_3K(Boc)-$(Boc)_2$NDI monomer (5.5) Upon completion of β-Ala_3K linker synthesis on solid phase, the terminal Fmoc was deprotected with 20% piperdine. The β-Ala_3K linker was coupled to compound **5.3** (3 eq) with the resin using PyBop (6 eq) and N-methylmorpholine (12 eq) in NMP. The mixture was shaken for 12 hours and drained. After draining, the resin was washed three times each with DMF, IPA, and DMF. The resin was cleaved using 30% hexafluoroisopropanol (HFIP) in DCM for 2 hours. The fully protected peptide was purified by preparative HPLC using a gradient of 33-35% B over 30 minutes. The product eluted at 13.5 minutes and lyophilization of pure fractions gave compound **5.5** as a white powder (100.3 mg, 19%). HRMS (ESI) calculated for $C_{49}H_{69}N_{10}O_{16}$ $(M+H)^+$: m/z 1053.489, found 1053.489.

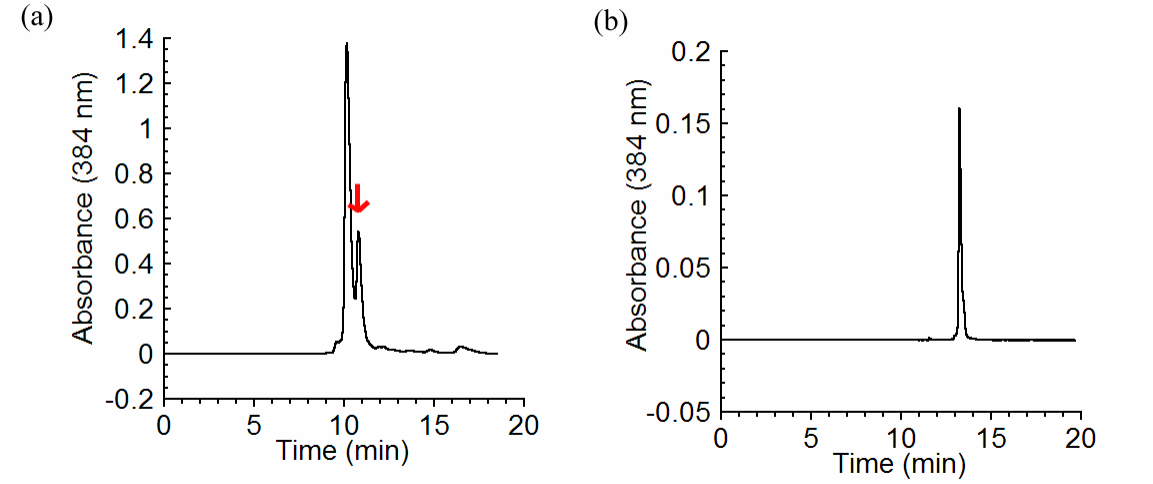


Figure 5.15 Compound **5.5** (a) crude and (b) pure HPLC. Red arrow indicates product in crude mixture.

Cyclic NDI bisintercalator (5.6) Compound **5.3** (4.36 mg, 6.8 x 10^{-6} mol) was dissolved in dry DMF (5 ml). The active esters were preformed by addition of pentafluorophenyl diphenylphosphinate (FDPP) (5.46 mg, 1.4 x 10^{-5} mol) and diisopropylethylamine (2.5 µl, 1.4 x 10^{-5} mol). The reaction was stirred for 1 hour. Compound **5.4** (10.1 mg, 6.8 x 10^{-}

^{6} mol) was dissolved in dry DMF (5 ml) and diisopropylethylamine (2.4 µl, 1.3 x 10^{-5} mol) and added to the reaction mixture. The reaction was stirred for 36 hours then concentrated under reduced pressure. The residue was dissolved in a 1:1:0.2 DCM/TFA/water mixture (10 ml) and stirred for 1 hour. The solvent was removed via a stream of nitrogen gas and the product was purified by preparative HPLC using a gradient of 10 to 14% over 40 minutes. The product eluted at 16 minutes and lyophilization of pure fractions gave compound **5.6** as a white powder (3.93 mg, 40%). HRMS (ESI) calculated for $C_{68}H_{85}N_{20}O_{18}$ $(M+H)^{+}$: m/z 1469.634, found 1469.630.

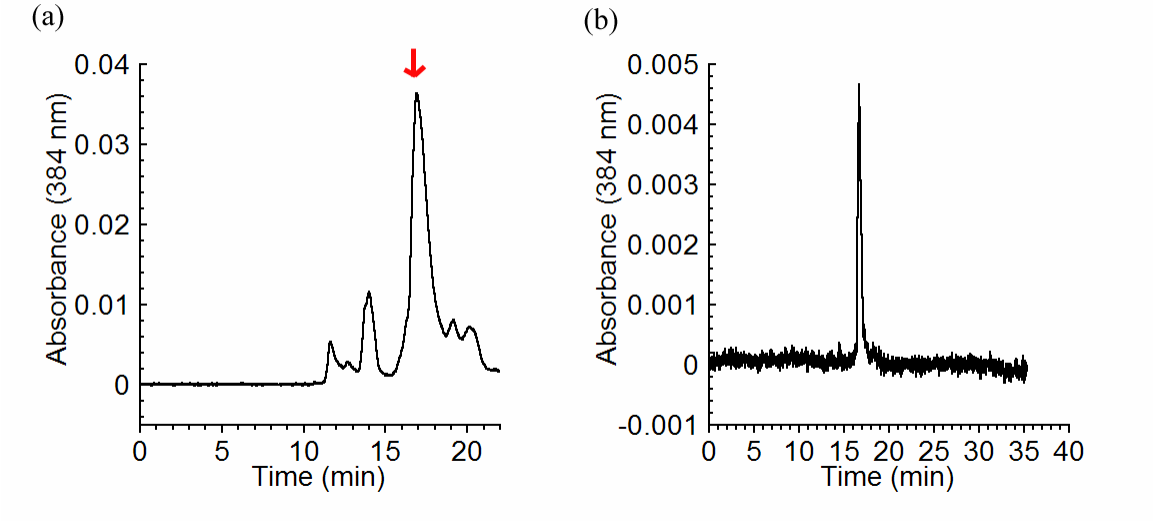


Figure 5.16 Compound **5.6** (a) crude and (b) pure HPLC. Red arrow indicates product in crude mixture.

Linear NDI bisintercalator (5.7) Compound **5.3** (6.28 mg, 9.8 x 10^{-6} mol) was dissolved in dry DMF (2.6 ml). The active esters were preformed by addition of PyBop (10.2 mg, 1.9 x 10^{-5} mol) and diisopropylethylamine (3.4 µl, 1.9 x 10^{-5} mol). The reaction was stirred for 1 hour. Compound **5.5** (5.2 mg, 4.9 x 10^{-6} mol) was dissolved in dry DMF (5 ml) and diisopropylethylamine (1.7 µl, 9.8 x 10^{-6} mol) and added to the reaction mixture. The reaction was stirred for 24 hours then concentrated under reduced pressure. The

residue was dissolved in a 1:1:0.2 DCM/TFA/water mixture (10 ml) and stirred for 1 hour. The solvent was removed via a stream of nitrogen gas and the product was purified by preparative HPLC using a gradient of 9 to 14% over 40 minutes. The product eluted at 35 minutes and lyophilization of pure fractions gave compound **5.6** as a white powder (1.14 mg, 20%). HRMS (CI) calculated for $C_{54}H_{59}N_{14}O_{17}$ $(M+H)^+$: m/z 1175.418, found 1175.417.

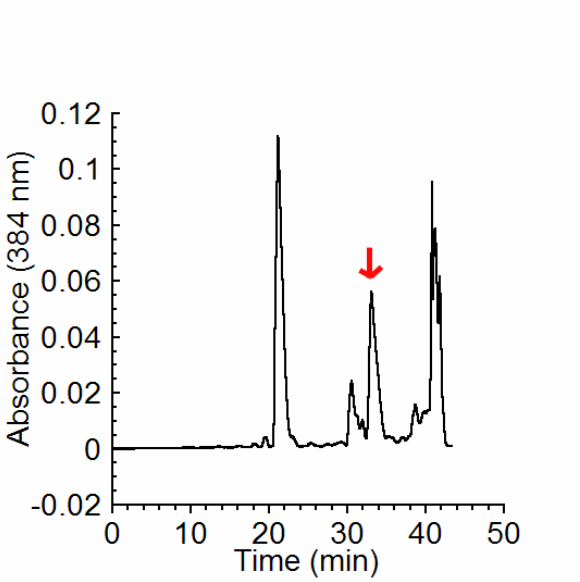


Figure 5.17 Crude HPLC of compound **5.7**. Red arrow indicates product in crude mixture.

Isothermal titration calorimetry All ITC titrations were performed in degassed 50 mM sodium phosphate buffer at pH 7. Concentrations of **5.8**, **5.6**, and **5.7** were determined by UV spectrophotometry (**5.6** and **5.7** λ = 364 nm, ε = 37,000, 5.8 λ = 296 nm, ε = 8,100). The cyclic NDI dimer **5.6** (1.4 ml) was placed in the cell at 0.6 mM concentration and the DAN monomer **5.8** (7.85 mM) was titrated into the cell with 35 injections of 7 µl each at 25° C. The injections were performed with a spacing of 180 seconds with the reference power set at 60. The linear NDI dimer **5.7** was titrated similarly except at a concentration of 0.89 mM with 8 µl injections of a 10.73 mM solution of DAN monomer **5.8**. Data was

fitted using assuming 1:1 binding without the first data point after subtraction of the near zero heat of dilution blank.

NMR titration NMR spectra were taken in 0.75 ml of 50 mM sodium phosphate buffer at pH 7 that had been lyophilized twice from D_2O using trimethylsilyl propionic acid-d_4 (TSP-d_4) as a reference. Concentrations of **5.9** and **5.10** stock solution were determined by UV (**5.9** λ = 384 nm, ε = 24,700, **5.10** λ = 296 nm, ε = 8,100). The final concentration of NDI monomer **5.9** in each of the 7 NMR samples was 0.375 mM with the DAN monomer **5.10** concentrations ranging from 0 to 3 mM for 0, 0.25, 0.5, 1, 2, 4, and 8 equivalents relative to **5.9**. The data was plotted as concentration of **5.10** versus the chemical shift of the NDI aromatic protons in ppm and fitted with assuming 1-to-1 binding.

References

Aggeli, A.; Bell, M.; Carrick, L.M.; Fishwick C.W.G.; Harding, R.; Mawer, P.J.; Radford, S.E.; Strong, A.E.; Boden, N. "pH as a trigger of peptide β-sheet self-assembly and reversible switching between nematic and isotropic phases" J. Am. Chem. Soc. (2003), 125, 9619.

Aggeli, A.; Nyrkova, I.A.; Bell, M.; Harding, R.; Carrick, L.; McLeish, T.C.B.; Semenov, A.N.; Boden, N. "Hierarchical self-assembly of chiral rod-like molecules as a model for peptide β-sheet tapes, ribbons, fibrils, and fibers" Proc. Natl. Acad. Sci. U.S.A. (2001), 98, 11857.

Amabilino, D.B.; Ashton, P.R.; Brown, C.L.; Cordova, E.; Godinez, L.A.; Goodnow, T.T.; Kaifer, A.E.; Newton, S.P.; Pietraszkiewicz, M.; Philp, D.; Raymo, F.M.; Reder, A.S.; Rutland, M.T.; Slawin, A.M.Z.; Spencer, N.; Stoddart, J.F.; Williams, D.J. "Molecular meccano: 2. Self-assembly of [n]catenanes" J. Am. Chem. Soc. (1995), 117, 1271.

Aravinda, S.; Harinia, V.V.; Shamala, N.; Das, C.; Balaram, P. "Stucture and assembly of designed β-hairpin peptides in crystals as models for β-sheet aggregation" Biochemistry (2004), 43, 1832.

Aravinda, S.; Shamala, N.; Chittaranjan, D.; Sriranjini, A.; Karle, I.L.; Balaram, P. "Aromatic-aromatic interaction in crystal structures of helical peptide scaffolds containing projecting phenylalanine residues" J. Am. Chem. Soc. (2003), 125, 5308.

Aravinda, S.; Shamala, N.; Raijkishore, R.; Gopi, H.N.; Balaram, P. "A crystalline β-hairpin peptide nucleated by a type I' Aib-D-Ala β-turn: Evidence for cross-strand aromatic interactions" Angew. Chem. Int. Ed. (2002), 41, 3863.

Ashton, P.R.; Phip, D.; Spencer, N.; Stoddart, J.F. "The self-assembly of [n]pseudorotaxanes" J. Chem. Soc., Chem. Commun. (1991), 23, 1677.

Augenstine, L.G.; Ghiron, C.A. "The inactivation of trypsin by ultraviolet light, I. The correlation of inactivation with the disruption of constituent cystine" Proc. Natl. Acad. Sci. U.S.A. (1961), 47, 1530.

Azriel, R.; Gazit, E. "Analysis of the minimal amyloid-forming fragment of the islet amyloid peptide" J. Biol. Chem. (2001), 276, 34156.

Balakrishnan, K.; Datar, A.; Zhang, W.; Yang, X.; Naddo, T.; Huang, J.; Zuo, J.; Yen, M.; Moore, J.S.; Zang, L. "Nanofibril self-assembly of an arylene ethynylene macrocycle" J. Am. Chem. Soc. (2006), 128, 6576.

Baldwin, J.E.; North, M.; Flinn, A. "A convenient new synthesis of homoserine lactone" Tetrahedron Lett. (1987), 28, 3167.

Barbas, C.F.; Burton, D.R.; Scott, J.K.; Silverman, G.J. "Phage Display: A Laboratory Manual." Cold Spring Harbor Press: Cold Spring Harbor, NY (2001).

Bassette, P.H.; Rice, J.J.; Daugherty, P.S. "Rapid isolation of high-affinity protein binding peptides using bacterial display" Protein Eng. Des. Sel. (2004), 17, 731.

Ben-Ishai, D.; Berger, A. "Cleavage of N-carbobenzoxy groups by dry hydrogen bromide and hydrogen chloride" J. Org. Chem. (1952), 17, 1564.

Binder, W.H.; Smrzka, O.W. "Self-assembly of fibers and fibrils" Angew. Chem. Int. Ed. (2006), 45, 7324.

Brunsveld, L.; Meijer, E.W.; Prince, R.B.; Moore, J.S. "Self-assembly of folded *m*-phenylene ethynylene oligomers into helical columns" J. Am. Chem. Soc. (2001), 123, 7978.

Burdick, J.A.; Stevens, M.M. "Biomaterials, Artificial Organs and Tissue Engineering" Hench, L.; Jones, J.R., Eds. (2005), Woodhead Publishing: Cambridge.

Burley, S.K.; Petsko, G.A. "Aromatic-aromatic interactions: A mechanism of protein structure stabilization" Science (1985), 229, 23.

Burritt, J.B.; Bond, C.W.; Doss, K.W.; Jesaitis, A.J. "Filamentous phage display of oligopeptide libraries" Anal. Biochem. (1996), 238, 1.

Butterfield, S.M.; Goodman, C.M.; Rotello, V.M.; Waters, M.L. "A peptide flavoprotein mimic: flavin recognition and redox potential modulation in water by a designed β hairpin" Angew. Chem. Int. Ed. (2004), 43, 724.

Butterfield, S.M.; Patel, P.R.; Waters, M.L. "Contribution of aromatic interactions to α-helix stability" J. Am. Chem. Soc. (2002), 124, 9751.

Chakraborty, C.; Nandi, S.; Jana, S. "Prion disease: A deadly disease for protein misfolding" Curr. Pharm. Biotechno. (2005), 6, 167.

Chen, F.M. . "Methods for the Studies of Drug Dissociation from DNA." Humana Press: Totowa, NJ (1997).

Chen, P. "Self-assembly of ionic-complementary peptides: A physicochemical viewpoint" Colloids Surf. A (2005), 261, 3.

Cherney, R.J.; Wang, L. "Efficient Mitsunobu reactionswith N-phenylfluorenyl or N-trityl serine esters" J. Org. Chem. (1996), 61, 2544.

Chiti, F.; Dobson, C.M. "Protein misfolding, functiona amyloid, and human disease" Annu. Rev. Biochem. (2006), 75, 333.

Chiti, F.; Stefani, M.; Taddel, N.; Ramponi, G.; Dobson, C.M. "Rationalization of the effects of mutations on peptide and protein aggregation rates" Nature (2003), 424, 805.

Chu, Y.; Sorey, S.; Hoffman, D.W.; Iverson, B.L. "Structural characterization of a rigidified threading bisintercalator" J. Am. Chem. Soc. (2007a), 129, 1304.

Chu, Y. "DNA threading intercalation: Building sequence-specific linear rigidified and cyclic bisintercalators" University of Texas at Austin, Ph.D. dissertation, 2007b.

Ciani, B.; Hutchinson, E.G.; Session, R.B.; Woolfson, D.N. "A designed system for assessing how sequence affects α to β conformational transitions in proteins" J. Biol. Chem. (2002), 277, 10150.

Classens, C.G.; Stoddart, J.F. "Review commentary: π-π interactions in self-assembly" J. Phys. Org. Chem. (1997), 10, 254.

Cochran, A.G.; Skelton, N.J.; Starovasnik, M.A. "Tryptophan zippers: Stable, monomeric β-hairpins" Proc. Natl. Acad. Sci. U.S.A. (2001), 98, 5578.

Collier, J.H.; Hu, B.H.; Ruberti, J.W.; Zhang, J.; Shum, P.; Thompson, D.H.; Messersmith, P.B. "Thermally and photochemically triggered self-assembly of peptide hydrogels" J. Am. Chem. Soc. (2001), 123, 9463.

Connors, K.A. "Binding Constants: The Measurement of Molecular Complex Stability." John Wiley and Sons: New York (1987).

Cookson, M.R. "The biochemistry of Parkinson's disease" Annu. Rev. Biochem. (2005), 74, 29.

Cubberly, M.S.; Iverson, B.L. "^{1}H NMR investigation of solvent effects in aromatic stacking interactions" J. Am. Chem. Soc. (2001), 123, 7560.

de Loos, M.; Feringa, B.L.; van Esch, J.H. "Design and application of self-assembled low molecular weight hydrogels" Eur. J. Org. Chem. (2005), 17, 3615.

Deechongkit, S.; Powers, E.T.; You, S.L.; Kelly, J.W. "Controlling the morphology of cross β-sheet assemblies by rational design" J. Am. Chem. Soc. (2005), 127, 8562.

Deming, T.J. "Polypeptide hydrogels via a unique assembly mechanism" Soft Matter. (2005), 1, 28.

Diamond, R. "Real-space refinement of the structure of hen egg-white lysozyme" J. Mol. Biol. (1974), 82, 371.

Dobson, C. M. "Protein folding and misfolding" Nature (2003), 426, 884.

Finder, V.H.; Glockshuber, R. "Amyloid-β aggregation" Neurodegenerative Dis. 2007, 4, 13

Fischer, R.; Mader, O.; Jung, G.; Brock, R. "Extending the applicability of carboxyfluorescein in solid-phase synthesis" Bioconjugate Chem. (2003), 14, 653.

Francisco, J.A.; Campbell, R.; Iverson, B.L.; Georgiou, G. "Production and fluorescence-activated cell sorting of Escherichia coli expressing a functional antibody fragment on the external surface" Proc. Natl. Acad. Sci. U.S.A. (1993), 90, 10444.

Fremont, D.H.; Hendrickson, W.A.; Marrack, P.; Kappler, J. "Structures of an MHC class II molecule with covalently bound single peptides" Science (1996), 272, 1001.

Gabriel, G.J.; Iverson, B.L. "Aromatic oligomers that form heteroduplexes in aqueous solutions" J. Am. Chem. Soc. (2002), 124, 15174.

Gabriel, G.J.; Sorey, S.; Iverson, B.L. "Altering the folding patterns of naphthyl trimers" J. Am. Chem. Soc. (2005), 127, 2637.

Gardinier, K.M.; Khoury, R.G.; Lehn, J.-M. "Enforced helicity: efficient access to self-organized helical molecular strands by the imine route" Chem. Eur. J. (2000), 6, 4124.

Gazit, E.A. "A possible role for π-stacking in the self-assembly of amyloid fibrils" FASEB J. (2002), 16, 77.

Ghosh, S.; Ramakrishnan, S. "Aromatic donor-acceptor charge-transfer and metal-ion-complexation-assisted folding of a synthetic polymer" Angew. Chem. Int. Ed. (2004), 43, 3264.

Ghosh, S.; Ramakrishnan, S. "Structural fine-tuning of (-donor-*spacer*-acceptor-*spacer*-)$_n$ type foldamers. Effect of spacer segment length, temperature, and metal-ion complexation on the folding process" Macromolecules (2005a), 38, 676.

Ghosh, S.; Ramakrishnan, S. "Small-molecule-induced folding of a synthetic polymer" Angew. Chem. Int. Ed. (2005b), 44, 5441.

Goodin, D.B.; McRee, D.E. "The Asp-His-Fe triad of dytochrome c peroxidase controls the reduction potential, electronic structure, and coupling of the tryptophan fFree radical to the Heme" (1993), 32, 3313.

Gregersen, N.; Bross, P.; Vang, S.; Christensen, J.H. "Protein misfolding and human disease" Annu. Rev. Genom. Hum. G. (2006), *7*, 103.

Guelev, V.; Lee, J.; Ward, J.; Sorey, S.; Hoffman, D.W.; Iverson, B.L. "Peptide bisintercalator binds DNA via threading mode with sequence specific contacts in the major groove" Chem. Biol. (2001), 8, 415.

Guelev, V.; Sorey, S.; Hoffman, D.W.; Iverson, B.L. "Changing DNA grooves-a 1,4,5,8-naphthalene tetracarboxylic diimide bis-intercalator with the linker $(\beta\text{-Ala})_3$Lys in the minor groove" J. Am. Chem. Soc. (2002), 124, 2864.

Guéron, M.; Leroy, J.L. Base-pair opening in double-stranded nucleic acids. "Nucleic Acids and Molecular Biology" Eckstein, F. and Lilley, D.M.J. (eds), Springer-Verlag: New York (1992), vol. 6, pp. 1-22.

Gupta, R.K. Polymer and Composite Rheology; Marcel Dekker Inc.: New York, 2000.

Haines-Butterick, L.; Rajagopal, K.; Branco, M.; Salick, D.; Rughani, R.; Pilarz, M.; Lamm, M.S.; Pochan, D.J.; Schneider, J.P. "Controlling hydrogelation kinetics by peptide design for three-dimensional encapsulation and injectable delivery of cells" Proc. Natl. Acad. Sci. U.S.A. (2007), 104, 7791.

Haldar, D.; Jiang, H.; Leger, J.M.; Huc, I. "Interstrand interactions between side chains in a double-helical foldamer" Angew. Chem. Int. Ed. (2006), 45, 5483.

Hamilton, D.G.; Sanders, J.K.M.; Davies, J.E.; Clegg, W.; Teat, S.J. "Neutral [2]catenanes from oxidative coupling of π-stacked components" Chem. Commun. (1997), 9, 897.

Hamley, I.W. "Peptide fibrillization" Angew. Chem. Int. Ed. (2007), 46, 8128.

Hart, D.S.; Gehrke, S.H. "Thermally associating polypeptides designed for drug delivery produced by genetically engineered cells" J. Pharm. Sci. (2007), 96, 484.

Harvey, B.R.; Georgiou, G.; Hayhurst, A.; Iverson, B.L.; Rogers, G.K. "Anchored periplasmic expression, a versatile technology for the isolation of high-affinity antibodies from Escherichia coli-expressed libraries" Proc. Natl. Acad. Sci. U.S.A. (2004), 101, 9193.

Hirata, A.; Sugimoto, K.; Konno, T.; Morii, T. "Amyloid-forming propensity of the hydrophobic non-natural amino acid on the fibril-forming core peptide of human tau" Bioorg. Med. Chem. Lett. (2007), 17, 2971.

Hoeppener, J.W.M.; Lips, C.J.M. "Role of islet amyloid in type 2 diabetes mellitus" Int. J. Biochem. Cell B. (2006), 38, 726.

Hofstädter, K.; Stuart, F.; Jiang, L.; Vrijbloed, J.W.; Robinson, J.A. "On the importance of being aromatic at an antibody-protein antigen interface: Mutagenesis of the extracellular interferon γ receptor and recognition by the neutralizing antibody A6" J. Mol. Biol. (1999), 285, 805.

Horne, W.S.; Ashkenasy, N.; Ghadiri, M.R. "Modulating charge transfer through cyclic D,L-α-peptide self-assembly" Chem. Eur. J. (2005),11, 1137.

Hughes, R.M.; Benshoff, M.L.; Waters, M.L. "Effects of chain length and N-methylation on a cation-π interaction in a β-hairpin peptide" Chem. Eur. J. (2007), 13, 5753.

Hughes, R.M.; Waters, M.L. "Effects of lysine acetylation in a β-hairpin peptide: Comparison of an amide-π and a cation-π interaction" J. Am. Chem. Soc. (2006), 128, 13586.

Hughes, R.M.; Waters, M.L. "Influence of N-methylation on a cation-π interaction produces a remarkably stable β-hairpin peptide" J. Am. Chem. Soc (2005), 127, 6518.

Hunter, C.A.; Lawson, K.R.; Perkins, J.; Urch, C.J. "Aromatic interactions" J. Chem. Soc. Perkin Trans. 2 (2001), 651.

Hunter, C.A.; Sanders, J.K.M. "The nature of π-π interactions" J. Am. Chem. Soc. (1990), 112, 5225.

Iijima, T.; Vignon, S.A.; Tseng, H.-R.; Jarrosson, T.; Sanders, J.K.M.; Marchioni, F.; Ventur, M.; Apostoli, E.; Balzani, V.; Stoddart, J.F. "Controllable donor-acceptor neutral [2]-rotaxanes" Chem. Eur. J. (2004), 10, 6375.

Invitrogen Inc. The Handbook- A Guide to Fluorescent Probes and Labeling Technologies. (Carlsbad, California, 92008)(2005) Tenth Edition.

Jourdan, M.; Garcia, J.; Lhomme, J.; Teulade-Fichou, M.P.; Vigneron, J.P.; Lehn, J.-M. "Threading bis-intercalation of a macrocyclic bisacridine at abasic sites in DNA: nuclear magnetic resonance and molecular modeling study" Biochemistry (1999), 38, 14205.

Kammerer, R.A.; Kostrewa, K.; Zurdo, J.; Detken, A.; Garcia-Echeverria, C.; Green, J.D.; Muller, S.A.; Meier, B.H.; Winkler, F.K.; Dobson, C.M.; Steinmetz, M.O. "Exploring amyloid formation by a *de novo* design" Proc. Natl. Acad. Sci. U.S.A. (2004), 101, 4435.

Karplus, P.A. "Hydrophobicity regained" Protein Sci. (1997), 6, 1302.

Kayed, R.; Head, E.; Thompson, J.L.; McIntire, T.M.; Milton, S.C.; Cotman, C.W.; Glabe, C.G. "Common structure of soluble amyloid oligomers implies common mechanism of pathogenesis" Science (2003), 300, 486.

Khan, A.; Kaiser, C.; Hecht, S. "Prototype of a photoswitchable foldamer" Angew. Chem. Int. Ed. (2006), 45, 1878.

Kiehna, S.E.; Laughrey, Z.R.; Waters, M.L. "Evaluation of carbohydrate-π interaction in a peptide model system" Chem. Commun. (2007), 4026.

Kovacs, J.M.; Mant, C.T.; Hodges, R.S. "Determination of intrinsic hydrophilicity/hydrophobicity of amino acid side chains in peptides in the absence of nearest-neighbor or conformational effects" Biopolymers (2006), 84, 283.

Krebber, A.; Bornhauser, S.; Burmester, J.; Honegger, A.; Willuda, J.; Bosshard, H.R.; Plückthun, A. "Reliable cloning of functional antibody variable domains from hybridomas and spleen cell repertoires employing a reengineered phage display system" J. Immun. Methods (1997), 201, 35.

Kryger, G.; Silman, I.; Sussman, J.L. "Three-dimensional structure of a complex of E2020 with acetylcholinesterase from *Torpedo californica*" J. Physiol-Paris (1998), 92, 191.

Kryger, G.; Silman, L.; Sussman, J.L. "Structure of acetylcholinesterase complexed with E2020 (Aricept®): implications for the design of new anti-Alzheimer drugs" Structure (1999), 7, 297.

Kuner, P.; Bohrmann, B.; Tjernberg, L.O.; Naslund, J.; Huber, G.; Celenk, S.; Gruninger-Leitch, F; Richard, J.G.; Kemp, J.A.; Nordsted, C. "Controlling polymerization of beta-amyloid and prion-derived peptides with synthetic small molecule ligands" J. Biol. Chem. (2000), 275, 1673.

Lahiri, S.; Thompson, J.L.; Moore, J.S. "Solvophobically driven π-stacking of phenylene ethynylene macrocycles and oligomers" J. Am. Chem. Soc. (2000), 122, 11315.

Landsteiner, K.; Jacobs, J. "Studies on the sensitization of animals with simple chemical compounds" J. Exp. Med. (1935), 61, 643.

Lee, J.; Guelev, V.; Sorey, S.; Hoffman, D.W.; Iverson, B.L. "NMR structural analysis of a modular threading tetraintercalator bound to DNA" J. Am. Chem. Soc. (2004), 126, 14036.

Litvinchuk, S.; Tanaka, H.; Miyatake, T.; Pasini, D.; Tanaka, T.; Bollot, G.; Mareda, J.; Matile, S. "Synthetic pores with reactive signal amplifiers as artificial tongues" Nat. Mater. (2007), 6, 576.

Liu, D.; Williamson, D.A.; Kennedy, M.L.; Williams, T.D.; Morton, M.M.; Benson, D.R. "Aromatic side chain-porphyrin interactions in designed hemoproteins" J. Am. Chem. Soc. (1999), 121, 11798.

Liu, S.; Zavalij, P.Y.; Lam, Y.F.; Isaacs, L. "Refolding foldamers: Triazene-arylene oligomers that change shape with chemical stimuli" J. Am. Chem. Soc. (2007), 129, 11232.

Lokey, R.S.; Iverson, B.L. "Synthetic molecules that fold into a pleated secondary structure in solution" Nature (1995), 375, 303.

Lorenzo, A.; Yankner, B.A. "Beta-amyloid neurotoxicity requires fibril formation and is inhibited by congo red" Proc. Natl. Acad. Sci. U.S.A. (1994), 91, 12243.

Ma, J.C.; Dougherty, D.A. "The cation-pi interaction" Chem. Rev. (1997), 97, 1303.

Macedo, P.; Marrano, C.; Keillor, J. W. "Synthesis of dipeptide-bound epoxides and α,β-unsaturated amides as potential irreversible transglutaminase inhibitors" Bioorg. Med. Chem. Lett. (2002), 10, 355.

Makin, O.S.; Atkins, E.; Sikorski, P.; Johansson, J.; Serpell, L.C. "Molecular basis for amyloid fibril formation and stability" Proc. Natl. Acad. Sci. U.S.A. (2005), 102, 315.

Malone, J.F.; Murray, C.M.; Charlton, M.H.; Docherty, R.; Lavery, A.J. "X-H-π interactions: Theoretical and crystallographic observations" J. Chem. Soc., Faraday Trans. (1997), 93, 3429.

Mao, C.; Kisaalita, W.S. "Characterization of 3-D collagen hydrogels for functional cell-based biosensing" Biosens. Bioelectron. (2004), 19, 1075.

Matsumura, S.; Uemura, S.; Mihara, H. "Fabrication of nanofibers with uniform morphology by self-assembly of designed peptides" Chem. Eur. J. (2004), 10, 2789.

McGaughey, G.B.; Gagne, M.; Rappe, A.K. "π-Stacking interactions" J. Biol. Chem. (1998), 273, 15458.

McGregor, D.P.; Mollogy, P.E.; Cunningham, C.; Harris, W.J. "Spontaneous assembly of bivalent single chain antibody fragments in Escherichia coli" Mol. Immunol. (1994), 31, 219.

Measey, T.J.; Schweitzer-Stenner, R. "Aggregation of the amphiphathic peptides $(AAKA)_n$ into antiparallel β-sheets" J. Am. Chem. Soc. (2006), 128, 13324.

Meinhardt, J.; Tartaglia, G.G.; Pawar, A.; Christopeit, T.; Hortschansky, P.; Schroeckh, V.; Dobson, C.M.; Verdruscolo, M.; Fandrich, M. "Similarities in the thermodynamics and kinetics of aggregation of disease-related Aβ(1-40) peptides" Protein Sci. (2007), 16, 1214.

Merlini, G.; Westermark, P. "The systemic amyloidoses: Clearer understanding of the molecular mechanisms offers hope for more effective therapies" J. Intern. Med. (2004), 255, 159.

Mesquida, P.; Riener, C.K.; MacPhee, C.E.; McKendry, R.A. "Morphology and mechanical stability of amyloid-like peptide fibrils" J. Mater. Sci. Matter M. (2007), 18, 1325.

Meyer, E.A.; Castellano, R.K.; Diederich, F. "Interactions with aromatic rings in chemical and biological recognition" Angew. Chem. Int. Ed. (2003), 42, 1210.

Mimna, R.; Camus, M.S.; Schmid, A.; Tuschscherer, G.; Lashuel, H.A.; Mutter, M. "Disruption of amyloid-derived peptide assemblies through the controlled induction of a β-sheet to α-helix transformation: Application of the switch concept" Angew. Chem. Int. Ed. (2007), 46, 2681.

Mingarro, I.; Blanquet, S.; Plateau, P. "Membrane-protein engineering" Trends Biotechnol. (1997), 15, 432.

Moonen, N.N.P.; Flood, A.H.; Fernandez, J.M.; Stoddart, J.F. "Towards a rational design of molecular switches and sensors from their basic building blocks" Top. Curr. Chem. (2005), 262, 99.

Morgan, R.S.; Tatsch, C.E.; Gushard, R.H.; McAdon, J.M.; Warme, P.K. "Chains of alternating sulfur and π-bonded atoms in eight small proteins" Int. J. Prot. Pept. Res. (1978), 11, 209.

Muraki, M. "The importance of CH/π interactions to the function of carbohydrate binding proteins" Protein Peptide Lett. (2002), 9, 195.

Muraki, M.; Harata, K. "Origin of carbohydrate recognition specificity of human lysozyme revealed by affinity labeling" (1996), 35, 13562.

Nelson, J.C.; Saven, J.G.; Moore, J.S.; Wolynes, P.G. "Solvophobically driven folding of nonbiological oligomers" Science (1997), 277, 1793.

Nemethy, G.; Scheraga, H.A. "Strong interaction between disulfide derivatives and aromatic groups in peptides and proteins" Biochem. Biophys. Res. Commun. (1981), 98, 482.

Nguyen, J.Q.; Iverson, B.L. "An amphiphilic folding molecule that undergoes an irreversible conformational change" J. Am. Chem. Soc. (1999), 121, 2639.

Northrop, B.H.; Braunschweig, A.B.; Mendes, P.M.; Dichtel, W.R.; Stoddard, J.F. "Molecular machines" Goddard, W.A. ed. Handbook of Nanoscience, Engineering, and Technology (2007), 2nd ed.

Nowak, A.P.; Breedveld, V.; Pakstis, L.; Ozbas, B.; Pine, D.J.; Pochan, D.; Deming, T.J. "Rapidly recovering hydrogel scaffolds from self-assembling diblock copolypeptide amphiphiles" Nature (2002), 417, 424.

Olsen, M.J.; Stephens, D.; Griffiths, D.; Daugherty, P.; Georgiou, G.; Iverson, B.L. "Function-based isolation of novel enzymes from a large library" Nat. Biotechnol. (2000), 18, 1071.

Orfao, A.; Ruiz-Argüelles, A. "General concepts about cell sorting techniques" Clinical Biochem. (1996), 29, 5.

Ozinskas, A.J.; Rosenthal, G.A. "Synthesis of L-canaline and γ-functional 2-aminobutyric acid derivatives" J. Org. Chem. (1986), 51, 5047.

Parameswaran, K.N. "A facile synthesis of biotin N-hydroxysuccimide ester" Org. Prep. Proced. Int. (1990), 22, 119.

Park, H.; Park, K. "Hydrogels in bioapplications" ACS Sympoisum Series (1996), 627, 2.

Park, K.; Shalaby, W.S.W.; Park, H. "Biodegradable Hydrogels for Drug Delivery" (1993), Technomic Publishing: Lancaster, PA.

Pengo, P.; Pantos, G.D.; Otto, S.; Sanders, J.K.M. "Efficient and mild microwave-assisted stepwise functionalization of naphthalenediimide with α-amino acids" J. Org. Chem. (2006), 71, 7063.

Pomerantz, W.C.; Abbott, N.L.; Gellman, S.H. "Lyotropic liquid crystals from designed helical β-peptides" J. Am. Chem. Soc. (2006), 128, 8730.

Ramachandran, S.; Trewhella, J.; Tseng, Y.; Yu, Y.B. "Coassembling peptide-based biomaterials: Effects of pairing equal and unequal chain length oligopeptides" Chem. Mater. (2006), 18, 6157.

Ray, S.; Drew, M.G.B.; Das, A.K.; Banerjee, A. "Supramolecular β-sheet and nanofibril formation by self-assembling tripeptides containing an N-terminally located γ-aminobutyric acid residue" Supramol. Chem. (2006), 18, 455.

Reczek, J.J. "Aromatic electron donor-acceptor interactions in noval supramolecular assemblies" University of Texas at Austin, Ph.D. dissertation, 2006.

Reczek, J.J.; Iverson, B.L. "Using aromatic donor acceptor interactions to affect macromolecular assembly" Macromolecules (2006a), 39, 5601.

Reczek, J.J.; Villazor, K.R.; Lynch, V.; Swager, T.M.; Iverson, B.L. "Tunable columnar mesophases utilizing C_2 symmetric aromatic donor-acceptor complexes" J. Am. Chem. Soc. (2006b), 128, 7995.

Reixach, N.; Deechongkit, S.; Jiang, X.; Kelly, J.W.; Buxbaum, J.N. "Tissue damage in the amyloidoses: Transthyretin monomers and nonnative oligomers are the major cytotoxic species in tissue culture" Proc. Natl. Acad. Sci. U.S.A. (2004), 101, 2817.

Ross, C.A. "Polyglutamine pathogenesis: Emergence of unifying mechanisms for Huntington's disease and related disorders" Neuron (2002), 35, 819.

Saha, S.; Flood, A.H.; Stoddart, J.F.; Impellizzeri, S.; Silvi, S.; Venturi, M.; Credi, A. "A redox-driven multicomponent molecular shuttle" J. Am. Chem. Soc. (2007), 129, 12159.

Sal-Man, N.; Gerber, D.; Bloch, I.; Shai, Y. "Specificty in transmembrane helix-helix interactions mediated by aromatic residues" J. Biol. Chem. (2007), 282, 19753.

Sánchez-Quesada, J.; Isler, M.P.; Ghadiri, M.R. "Modulation ion channel properties of transmembrane peptide nanotubes through heteromeric supramolecular assemblies" J. Am. Chem. Soc. (2002), 124, 10004.

Sánchez-Quesada, J.; Kim, H.S.; Ghadiri, M.R. "A synthetic pore-mediated transmembrane transport of glutamic acid" Angew. Chem. Int. Ed. (2001), 40, 2503.

Schatz, P.J.; Cull, M.G.; Martin, E.L.; Gates, C.M. "Screening of peptide libraries linked to lac repressor" Methods in Enzymology (1996) 267(Combinatorial Chemistry), 171.

Schneider, J.P.; Pochan, D.J.; Ozbas, B.; Rajagopal, K.; Pakstis, L; Kretsinger, J. "Responsive hydrogels from the intramolecular folding and self-assembly of a designed peptide" J. Am. Chem. Soc. (2002), 124, 15030.

Seebach, D.; Hook, D.F.; Glattli, A. "Helices and other secondary structures of β- and γ-peptides" Biopolymers (2006), 84, 23.

Serrano, L.; Bycroft, M.; Fersht, A.R. "Aromatic-aromatic interactions and protein stability: Investigation by double-mutant cycles" J. Mol. Biol. (1991), 218, 465.

Shen, W.; Kornfield, J.A.; Tirrell, D.A. "Dynamic properties of artificial protein hydrogels assembled through aggregation of leucine aipper peptide domains" Macromolecules (2007), 40, 689.

Shin, S.B.Y.; Kirshenbaum, K. "Conformational rearrangements by water-soluble peptoid foldamers" Org. Lett. (2007), 9, 5003.

Stanger, H.E.; Gellman, S.H. "Rules for antiparallel α-Sheet design: D-Pro-Gly is superior to L-Asn-Gly for α-hairpin nucleation" J. Am. Chem. Soc. (2002), 124, 9372.

Steiner, T.; Koellner, G. "Hydrogen bonds with pi-acceptors in proteins: frequencies and role in stabilizing local 3D structures" J. Mol. Biol. (2001), 305, 535.

Stone, M.T.; Moore, J.S. "A water-soluble m-phenylene ethynylene foldamer" Org. Lett. (2004), 6, 469.

Tanaka, J.; Bollot, G.; Mareda, J.; Litvinchuk, S.; Tran, D.J.; Sakai, N.; Matile, S. "Synthetic pores with sticky π-clamps" Org. Biomol. Chem. (2007), 5, 1369.

Tanatani, A.; Mio, M.J.; Moore, J.S. "Chain length-dependent affinity of helical foldamers for a rodlike guest" J. Am. Chem. Soc. (2001), 123, 1792.

Tartaglia, G.G.; Cavalli, A.; Pellarin, R.; Caflisch, A. "The role of aromaticity, exposed surface, and dipole moment in determining protein aggregation rates" Protein Sci. (2004), 13, 1939.

Tatko, C.D.; Waters, M.L. "Comparison of C-H-π and hydrophobic interactions in a β-hairpin peptide: Impact on stability and specificity" J. Am. Chem. Soc. (2004), 126, 2028.

Tatko, C.D.; Waters, M.L. "Selective aromatic interactions in β-hairpin peptides" J. Am. Chem. Soc. (2002), 124. 9372.

Taylor, B.M.; Sarver, R.W.; Fici, G.; Poorman, R.A.; Lutzke, B.S.; Molinari, A.; Kawabe, T.; Kappenman, K.; Buhl, A.E.; Epps, D.E. "Spontaneous aggregation and cytotoxicity of the β-amyloid Aβ(1-40): A kinetic model" J. Protein Chem. (2003), 22, 31.

Teplow, D.B. “Structural and kinetic features of amyloid beta-protein fibrillogenesis” Amyloid (1998), 5, 121.

Terech, P. “Fibers and wires in organogels from low-mass compounds: typical structural and rheological properties” Ber. Busen-Ges. Phys. Chem. Chem. Phys. (1998), 102, 1630.

Terpe, K. “Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems” App. Microbiol. Biotechnol. (2003), 60, 523.

Toledano, S.; Williams, R.J.; Vineetha, J.; Ulijn, R. “Enzyme-triggered self-assembly of peptide hydrogels via reversed hydrolysis” J. Am. Chem. Soc. (2006), 128, 1070.

Tossi, A.; Sandri, L.; Giangaspero, A. “New consensus hydrophobicity scale extended to non-proteinogenic amino acids” Peptides (2002), 27, 416.

Twyman, L.J.; Allsop, D. “A short synthesis of the β-amyloid (Aβ) aggregation inhibitor 2-p-toluoyl-2-[4’-(3-diethylaminopropoxy)-phenyl]-benzofuran” Tetrahedron Lett. (1999), 40, 9383.

Um, S.H.; Lee, J.B.; Park, J.; Kwon, S.Y.; Umbach, C.C.; Luo, D. “Enzyme-catalyzed assembly of DNA hydrogel” Nat. Mater. (2006), 5, 797.

Veal, J.M.; Li, Y.; Zimmerman, S.C.; Lamberson, C.R.; Cory, M.; Zon, G.; Wilson, W.D. “Interaction of a macrocyclic bisacridine with DNA” Biochemistry (1990), 29, 10918.

Vignon, S.A.; Jarrosson, T.; Iijima, T.; Tseng, H.-R.; Sanders, J.K.M.; Stoddart, J.F. “Switchable neutral bistable rotaxanes” J. Am. Chem. Soc. (2004), 126, 9884.

Wang, K.; Keasling, J.D.; Muller, S.J. “Effects of the sequence and size of non-polar residues on self-assembly of amphiphilic peptides” Int. J. Biol. Macromol. (2005), 36, 232.

Waters, M.L. “Aromatic interaction in peptides: Impact on structure and function” Biopolymers (2004), 76, 435.

Wegner, G.J.; Lee, H.J.; Corn, R.M. “Characterization and optimization of peptide arrays for the study of epitope-antibody interactions using surface plasmon resonance imaging” Anal. Chem. (2002), 74, 5161.

Weissbluth, M. “Hypochromism” Q. Rev. Biophys. (1971), 4, 1.

Wilchek, M.; Bayer, E.A. “Applications of avidin-biotin technology: literature survey” Methods in Enzymology (1990), 184(Avidin-Biotin Technology), 14.

Wilcox, C.S. “Frontiers in Supramolecular Organic Chemistry and Photochemistry.” Schnieder, H.J. and Dürr, H., (eds), VCH: Weinheim (1991), pp 122-143.

Wu, C.W.; Sanborn, T.J.; Huang, K.; Zuckermann, R.N.; Barron, A.E. “Peptoid oligomers with α-chiral, aromatic side chains: Sequence requirements for the formation of stable peptoid helices” J. Am. Chem. Soc. (2001), 123, 6778.

Xu, C.; Kopecek, J. “Self-assembling hydrogels” Polym. Bull. (2007), 58, 53.

Yamaguchi, K.; Yu, F.; Inouye, M. “A single amino acid determinant of the membrane localization of lipoproteins in E. coli” Cell (1988), 53, 423.

Yan, H. Saiani, A.; Gough, J.E.; Miller, A.F. “Thermoreversible protein hydrogel as cell scaffold” Biomacromolecules (2006), 7, 2776.

Yang, Z.; Ho, P.L.; Liang, G.; Chow, K.H.; Wang, Q.; Cao, Y.; Guo, Z.; Xu, B. “Using β-lactamase to trigger supramolecular hydrogelation” J. Am. Chem. Soc. (2007), 129, 266.

Yang, Z.; Liang, G.; Wang, L.; Xu, B. “Using a kinase/phosphotase switch to regulate a supramolecular hydrogel and forming the supramolecular hydrogel in vivo” J. Am. Chem. Soc. (2006), 128, 3038.

Yoo, B.; Pagel, M.D. “Peptidyl molecular imaging contrast agents using a new solid-phase peptide synthesis approach” Bioconjugate Chem. (2007), 18, 903.

Zhang, L.; Kauffman, G.S.; Pesti, J.A.; Yin, J. “Rearrangement of N_{α}-protected L-asparagines with iodosobenzene diacetate. A practical route to β-amino-L-alanine derivatives” J. Org. Chem. (1997), 62, 6918.

Zhang, W.; Horoszewski, D.; Decatur, J.; Nuckolls, C. “A folded, secondary structure in step-growth oligomers from covalently linked, crowded aromatics” J. Am. Chem. Soc. (2003), 125, 4870.

Zhang, Y.; Gu, H.; Yang, Z.; Xu, B. “Supramolecular hydrogels respond to ligand-receptor interaction” J. Am. Chem. Soc. (2003), 125, 13680.

Zhao, X.; Jia, M.-X.; Jiang, X.-K.; Wu, L.-Z.; Lia, Z.-T.; Chen, G.-J. “Zipper-featured δ peptide foldamers driven by donor-acceptor interactions. Design, synthesis, and characterization” J. Org. Chem. (2004), 69, 270.

Zhou, Q.-Z.; Jia, M.-X.; Shao, X.-B.; Wu, L.-Z.; Jiang, X.-K.; Li, Z.-T.; Chen, G.-J. “Self-assembly of a novel series of hetero-duplexes driven by donor-acceptor interaction” Tetrahedron (2005), 61, 7117.

Zhou, Q.-Z.; Jiang, X.-K.; Shao, X.-B.; Chen, G.-J.; Jia, M.-X.; Li, Z.-T. "First zipper-featured molecular duplexes driven by cooperative donor-acceptor interaction" Org. Lett. (2003), 5, 1955.

Zimmerman, S.C.; Lamberson, C.R.; Cory, M.; Fairley, T.A. "Topologically constrained bifunctional intercalators: DNA intercalation by a macrocyclic bisacridine" J. Am. Chem. Soc. (1989), 111, 6805.

Zych, A.J.; Iverson, B.L. "Conformational modularity of an abiotic secondary-structure motif in aqueous solution" Helv. Chim. Acta. (2002), 85, 3294.

Zych, A.J.; Iverson, B.L. "Synthesis and conformational characterization of tethered, self-complexing 1,5-dialkoxynaphthalene/1,4,5,8-naphthalenecarboxylic diimide systems" J. Am. Chem. Soc. (2000), 122, 8898.

Zych, A.J.; Iverson, B.L."Synthesis and conformational characterization of tethered, self-comlexing 1,5-dialkoxynaphthalene/1,4,5,8-naphthalenetetracarboxylic diimide systems" J. Am. Chem. Soc. (2000), 122, 8898.

Vita

Valerie Jean Bradford was born the oldest of two children of Larry Dzubeck and Joyce O'Neill on April 27, 1980 in the city of Teaneck, NJ. She graduated from Parsippany Hills High school in Parsippany, NJ in 1998 and attended the University of Delaware in the fall of that year. She began her research in bioorganic chemistry in the winter of 2000 as an undergraduate researcher in the lab of Joel P. Schneider where she worked until the summer of 2002. In the summer of 2001, she performed research as an intern at Eli Lilly and Company in the lead generation chemistry group. She received her Honors Bachelor of Science Degree with Distinction from the University of Delaware in 2002. She entered the Graduate School at The University of Texas at Austin in the fall of 2002.

Permanent address: 31 Oak Ln., Boonton, NJ 07005

This dissertation was typed by the author.